VOLUME 1. NO. 1 NOVEMBER, 1958

EDUCATIONAL RESEARCH

A Review for
Teachers and all concerned with
Progress in Education

SECRETARY'S OFFICE COPY
DO NOT REMOVE

Published for
THE NATIONAL FOUNDATION FOR EDUCATIONAL RESEARCH IN ENGLAND AND WALES
by
NEWNES EDUCATIONAL PUBLISHING CO. LTD.
Tower House, Southampton Street, Strand, London, W.C.2

EDUCATIONAL RESEARCH

Editor:
W. D. WALL
assisted by
ALFRED YATES D. A. PIDGEON
D. W. BELL

Educational Research is published for the National Foundation by Newnes Educational Publishing Co., Ltd., once a term, in November, February and June. The price per issue (post free) throughout the world is 5/6 and the annual subscription (post free) is 16/6.

Subscriptions, orders and advertisements should be sent to Newnes Educational Publishing Co., Tower House, Southampton Street, Strand, London, W.C.2.

Communications on general editorial matters including MSS. for publication, books and other printed matter for review or abstracting should be sent to the Director, N.F.E.R., 79, Wimpole Street, London, W.1.

Note to contributors: Contributions which make a readable synthesis of valid research on the practical problems posed by the English educational system will be particularly welcome. The length of such contributions should be between 4,000-7,000 words, though shorter notes will be considered. Abstracts of theses or similar material should be of sufficient length and detail to acquaint the reader with the main outcomes of the work, the sampling on which it has been based and the experimental statistical and other techniques used. Such synopses should be submitted in agreement with the Head of the University Department concerned.

All papers submitted for publication should conform to the conventional usage of the review and especially to the form in which bibliographical references are given. They should be typed (double space) on one side of the paper.

Intending contributors would be well advised in the first instance to send to the editor a brief synopsis of their paper with the elements of a bibliography.

In certain cases it may be possible for the Foundation's Information Service to aid in the gathering of relevant research material.

Only material which has not been published already elsewhere can normally be accepted for publication. The copyright in any material published is retained by the Editor. Reproduction in whole or in part is forbidden except (i) under the terms of the Fair Copying Declaration of the Royal Society or (ii) when specific permission has been given by the Editor.

Contributors receive free 15 copies of the number in which their paper appears. If they require further copies they should apply to the Editor at the time of returning proofs.

The conclusions drawn and the opinions expressed in papers are those of the authors. They should not be taken to represent the view of the N.F.E.R.

VOLUME 1. NO.1 NOVEMBER, 1958

EDUCATIONAL RESEARCH

A Review for
Teachers and all concerned with
Progress in Education

Published for
THE NATIONAL FOUNDATION FOR EDUCATIONAL
RESEARCH IN ENGLAND AND WALES

by
NEWNES EDUCATIONAL PUBLISHING CO. LTD.
Tower House, Southampton Street, Strand, London, W.C.2

EDITORIAL

EDITORIALS will not be a frequent or regularly recurring feature of this review but the first issue is evidently a special occasion. It seems appropriate to say something of the intentions in the minds of those responsible, something of the purposes we hope will be served and something of future editorial policy.

Many people in the education service are aware of the increasing gap between the worlds of research and the schools. With the best will in the world it is difficult for teachers to keep abreast of the new knowledge which is daily acquired by research workers in universities and in organizations like our own both in this country and abroad. The sheer volume of published work is more than one person can read and digest. Moreover, refinements in statistical and experimental methods, essential factors in the soundness of the work done, inevitably increase the difficulty of understanding for those who are not themselves professional research workers.

It is our purpose to bridge this gap. The articles in each number of the review will be substantial summaries of work in various research fields. We have no intention of 'popularizing' in the sense that articles are written down, or that they avoid such technical terminology as is necessary. The papers are intended to have the nature of comprehensive and authoritative statements of the principal facts and conclusions derived from research and are intended to provide our professional colleagues with the basis on which they can take practical decisions.

Certain topics are of such broad importance that there is much information available —more than can reasonably be digested into one paper. Such topics will, therefore, be treated in serial fashion, under a general heading such as *The Teaching of Reading* and the aim is that over a number of issues of the review a thorough study will be made. Such series of articles are planned to deal with practical problems of teaching method and organization. Single articles are in general devoted to more general problems—such as that by Professor Vernon in this issue.

In addition to the articles themselves, the review will have two other continuing features. The first of these is the abstract section in which from time to time major pieces of research work or research papers from foreign journals which may not be readily accessible in this country, will be fully summarized. The second is the bibliographical material which will either appear as an appendix to each article or in the form of a select annotated bibliography of the kind published in this issue.

We believe that a review of this kind is urgently needed by all those concerned with education; and we think that the proposed style and policy are approximately right. However, *Educational Research* is intended as a tool for those who read it and we shall be glad to have criticisms, suggestions for articles or indeed any comment which will enable us to give a better service.

W.D.W.

A NEW LOOK AT INTELLIGENCE TESTING

PHILIP E. VERNON

(*Professor of Educational Psychology, University of London, Institute of Education*).

Introduction.

TEACHERS who have followed the controversies over the past few years in the National Foundation's *Bulletin*, or in the periodical *Education*, and the *Times Educational Supplement*, may be pardoned if they have reached the conclusion that psychologists don't know their own minds. Apparently many of the fundamental principles of intelligence and intelligence testing which they imbibed from Ballard, Nunn, Hughes or Schonell are nowadays suspect, although still stoutly defended by a few authorities such as Professor Burt.

It might almost seem as though the whole raison d'être of using intelligence tests in educational selection and guidance had been demolished. That conclusion—I hope to show—is certainly not justified, though I would agree that many facts have emerged from research over the past 30 years or so which do indicate the need for some re-orientation in our theories of intelligence and in our practical application of tests. I should also claim that this 'new look' in intelligence testing represents a change of emphasis rather than a revolution; for actually much of what follows is in accord with what Burt, Schonell and I myself were teaching in the 1920's and 30's.

At the same time I would admit that psychologists are somewhat to blame for allowing a rather over-simplified and over-optimistic view of intelligence testing to become absorbed by student-teachers and by educationists generally—that is by people who could hardly be expected to keep in mind all the statistical snags and technical qualifications of which psychologists themselves should be aware. What are the essentials of this view which can no longer be accepted? They can be summarised as follows :—

The I.Q. as measured by standard intelligence tests, applied in a standard manner, measures the child's innate intellectual capacity which is determined by his genes. This general ability or *g* is the main factor underlying achievement in any direction, particularly in the educational sphere. Moreover, since this index of intelligence is very little affected by upbringing or environment, it enables us to predict, even in the junior school, the child's present and ultimate capacity for acquiring school learning and his suitability for a career involving high (or low) intelligence. If he is backward educationally and his I.Q. is low, we need seek no further explanation. But if his I.Q. is superior to his attainment, there are probably environmental or temperamental causes, or defects in schooling, which should be remedied so that he can work 'up to capacity'.

Variations In, and Diverse Kinds of, Intelligence.

The first aspect of this view which I must contest is the implicit assumption that intelligence is a single, clear-cut mental power or faculty, or in other words that any

one child will get the same I.Q. whatever test he is submitted to. In fact it is only partly true that different intelligence tests measure the same thing. They do overlap or correlate fairly highly with one another, but somewhat different results will certainly be obtained from, say, Moray House and National Foundation tests, more different from a verbal group test and individual Terman-Merrill, and still more different from verbal tests and non-verbal ones based on pictures, diagrams, or practical materials. Unfortunately it is far too common for teachers to apply one test and to enter the results under I.Q. on the record card, as though this was fixed immutably, regardless of the test used or the age of testing.

Now most children, when tested several times, do tend to score in the same band or region of I.Q. The average change when the tests are similar in character should be about 6 points, though this may rise to 10 points if the tests are more dissimilar, or if the period of time amounts to several years. Some are still more stable, but others are much more variable, and occasional gains or losses of 30 or even 40 points do occur. Suppose we regard 15 points as a big change, then about 1 in 9 will alter this much or more over a month or so with similar tests, and 1 in 6 will do so over several years or with dissimilar tests*.

A good deal of this variability is fortuitous : one test may have been given under noisy or otherwise unsuitable conditions ; the tester may add up the marks wrongly ; the children may have had some coaching ; the test norms or standards may be inaccurate ; or the spread of I.Q's—as measured by the standard deviation—may be abnormal ; any of these may differ between one test and another, and therefore upset the results to some extent. But even allowing for these sources of unreliability, it is clear that tests composed of different materials—verbal, numerical, spatial, etc.—or tests involving different functions such as inductive or deductive reasoning, memorising, fluency, etc. or using different sorts of items—multiple-choice vs. creative-response, speeded vs. unspeeded—all these measure somewhat different abilities.

Spearman was entirely justified in his claim that a common underlying factor *g* runs through all our mental abilities, and that this *g* is most strongly present in the higher thinking processes such as comprehension, reasoning, abstracting and generalising, or grasping relations. But he did not, like Burt and like the American factorists such as Thurstone, sufficiently realise the diversity as well as the unitariness of abilities.

This means then that intelligence is best regarded as a fluid collection of infinitely varied thinking abilities, and *g* is simply an average of whatever abilities the tester likes to include ; it cannot be pinned down to any specified intellectual function. Thus different tests or test batteries will inevitably yield somewhat different results, although they have much in common. There is no means of deciding that one test which samples one particular set of abilities is a better test of a hypothetical entity—intelligence—than is another test which samples a somewhat different set of abilities.

A further implication that many psychologists have drawn is that it would be preferable to give up the term intelligence, and instead build up batteries of tests which will best predict the particular kinds of intelligence in which we happen to be interested.

*These figures assume correlations of 0.90 and 0.70 respectively, also that the tests compared are accurately standardised with the same standard deviation of 15. If the norms differ, large variations may be more frequent.

For example we would have a set of Scholastic Aptitude tests for predicting educational success in predominantly linguistic courses, and a partly overlapping, partly different, set for predicting technical aptitude, and yet another for assessing brightness and quickness-in-the-uptake in everyday life, and so on. It is noticeable that the National Foundation issues tests called Verbal or Non-Verbal Tests rather than intelligence tests ; and Moray House has likewise recently transformed its series into Verbal Reasoning Tests.

In the scholastic field, there is no need to confine ourselves to the rather stereotyped and narrow range of types of item conventionally included in intelligence tests. Better predictions might be obtained from tests aimed at different kinds of reasoning, flexibility of mind, critical judgment, original thinking, and other partially distinguishable intellectual faculties.

Environmental Effects On Intelligence.

Let us now turn to changes in children's intelligence as they grow older. Such changes, attributable to environmental influences or to developments in the children's personalities, certainly do occur ; though they are probably less frequent and less marked than would appear merely from observing successive sets of I.Q's. because discrepancies so often arise—as we have seen—either from variations in the factorial content of the tests or from various fortuitous circumstances. However there is sufficient evidence of the effects of social class, upbringing and education to negate the view that intelligence is wholly inborn and therefore unalterable.

First, though, we should distinguish between improvements due to environmental factors and I.Q. gains brought about by coaching or practice on intelligence tests. These latter gains are certainly larger than we used to suppose : they range from about 3 to 12 points depending on the type of training given and the degree of previous sophistication of the children. But their most noteworthy features are that they are brought about very rapidly by 3 hours or less of training (thereafter there is hardly any further consistent improvement), and that they are highly specific to the particular types of items and the particular form of testing. There is no general spread to other kinds of intellectual performance. The children would seem merely to have acquired a familiarity with the mechanics of the tests on which they have been coached or practised. Thus, far from regarding this improvement as training their intelligence, I would suggest that most of the coaching that goes on for ' the 11 plus " is antithetical to true growth of their all-round thinking capacities.

Extensive investigations have been carried out into the effects of good and poor foster homes on the intelligence of orphan children ; also, identical twins, who have precisely the same heredity but who have been separated and brought up in different environments, have been studied. These researches rather uniformly indicate that the difference between a good and a poor home may raise or lower the I.Q. by about 10 points, or in more extreme cases by 20 points. At the same time there is no justification for those left-wing views which try to ascribe all intelligence differences to social class and upbringing, since some studies show small but significant correlations remaining between the intelligence of orphans and that of their true parents who have *not* brought

them up. Moreover it would be practically impossible to explain why children in the same family often show quite a wide range of I.Q., why professors sometimes have dull children and manual labourers bright ones, nor why orphans in a highly standardised institution do not all tend to approximate to the same dead level of intelligence, if environment were supreme. Whereas on genetic theory we would expect each brother and sister, or each orphan, to inherit a somewhat different proportion of the genes that make for good intelligence.

From twin data Burt and others have calculated that something like four-fifths of the differences among primary school children in intelligence should be attributed to hereditary causes. I do not regard that estimate as unreasonable when we consider the remarkable degree of uniformity of environment to which children are exposed up to 11 years. Certainly homes vary considerably in material circumstances and cultural stimulus. But so long as children are building up the simpler thinking capacities it is likely that most home and leisure environments provide the necessary exercise for their maturing brain processes. Indeed from 5 to 11 they are subjected to a highly standardised intellectual training at school.

But at the age when more complex and abstract thinking is developing, environments begin to differ more widely : they go to schools which exert very different degrees of intellectual pressure. By 15, half or more cease any formal education, enter largely routine jobs and indulge in leisure pursuits which make little call on intelligence ; whereas a privileged minority continue to exercise their minds to the utmost at school or university, are more likely to enter occupations that exercise their minds and to carry on cultural pursuits that make them think.

Hence I was able to show, by testing all the boys in a large city at 14 years that, after allowing for their different levels at 11+, those in the best secondary schools had gained 12 points over those in the weakest schools. The average modern school pupil had lagged 7 points behind the average grammar or technical school pupil.* Naturally, much of this difference might be due to the homes the different kinds of pupils came from rather than to the goodness of schooling as such; but it was, none the less, an environmental effect.

Similarly, Lorge in America and Husén in Sweden have tested the same individuals as children and as adults, and have shown that the later scores are definitely affected by the amount of schooling these individuals have received in the intervening period. Boys of the same I.Q. at the start, who had full secondary and further education, end up some 12 points ahead of those who left school at 14 to 15 and got no more education. Yet another piece of confirmatory evidence came from my analysis of the Progressive Matrices test scores of naval recruits of various ages during the war. This showed that the average capacity of adults to reason at this kind of test tends to decline with age; however, it stays up longer among men coming from skilled civilian occupations than among those whose civilian jobs appear to provide less intellectual stimulus.

*Yates and Pidgeon state that they are unable to confirm this finding. But they did not allow for the normal so-called regression effect. Had the grammar and modern schools in their follow-up had no effects on pupils' intelligence, the difference between their groups should have been 19.63 I.Q. points, if we guess the regression coefficient at 0.80. Since the obtained difference was 26.53 points, this indicates an alteration ,attributable to environment, of 7.1 points.

Such findings have tremendously important educational implications, since they show that full realisation of intellectual powers during the 'teens is often inhibited by early cessation of education, or by providing adolescents with a routine sort of schooling and conditions of work and leisure which deaden any desire to develop their minds further. While I personally favour education at the independent or maintained grammar school as doing most to draw out the intellectual faculties of the bright adolescent, I also feel a great deal of sympathy with the comprehensive school as providing more opportunities for the development of the pupil whose potentialities have been missed at 11+. The recent reports by the National Foundation and the British Psychological Society both show that the accuracy of selection at the end of primary schooling is remarkably high; in other words, that intelligence and attainment levels do remain fairly stable in the great majority. But they also agree that changes are to be expected in the ensuing years, and that therefore any rigid system of streaming or selection is bound to do injustice to at least 5 per cent. of children.

Much the same situation holds at the bottom end of the scale. A few per cent. of pupils drop so far behindhand in the primary school that it is to everyone's advantage to transfer them to special classes or to E.S.N. schools. But, partly because of the dogma of the fixity of the I.Q., it has been assumed that such pupils are incurably defective. They have become stuck in special schools and institutions. Indeed, many imbeciles have vegetated for the whole of their lives in mental hospitals, because we have failed to recognise that considerable improvements can occur in some—though certainly not in all —cases.

Recently, British psychologists, such as O'Connor, Clarke and Mundy, have been able to demonstrate significant I.Q. rises among feeble-minded patients, whose manual and other special abilities are utilised by training for appropriate jobs and who, by becoming responsible, largely self-supporting citizens escape from the deadening effects of the institutional environment. There is an important lesson here for the organisation of E.S.N. education. It is probable that a less rigid system of special classes might enable a large proportion to return to normal schooling who are at present left behind in segregated schools.

At the same time we must try to retain a balanced judgment. Not only are environmental effects on intellectual functioning rather limited in extent, but also they are by no means easy to control. We have no recipe for raising intelligence at will. We merely know, rather vaguely, that a lot of schooling is better than a little, that some schools are probably more effective than others, that cultured middle-class homes usually provide somewhat fuller realisation of potentialities than overcrowded slum homes. However, it should be possible through research to find out just what kinds of schooling, and what factors in the home, do most to stimulate all-round intellectual development.

The need for caution is also indicated by the negative results of so many experiments on transfer of training, which are quite rightly harped on in educational psychology textbooks. The 'new look' that I am advocating does not justify a return to the old belief in formal discipline, and the pre-eminence of a classical education. One other pointer that does emerge from investigations of pre-school children is the vital part played in intellectual development by emotional factors: there is strong evidence that the young

child who is emotionally secure in a democratic family environment is better able to exploit and develop his intellectual capacities than the child who is either over-protected or who feels rejected. And it may turn out that some subtle qualities of the classroom climate and the teacher's personality contribute as much, or more, to the intelligence of primary and secondary pupils as do the particular topics taught or the techniques of teaching.

The Psychology of Intellectual Development

There is another strong reason for revising our views on intelligence, namely, the intensive studies of children's psychological development that have been carried out by Gesell and his followers in America, and particularly by Piaget in Switzerland. Piaget shows that none of our mental capacities is innately given, or develops in isolation: even the baby's recognition of his rattle as a distinct object which can be moved by him or others and which makes noises, but nevertheless retains its identity, has to be built up through continuous exploration of, and experiment with, his environment. These early sensory-motor and perceptual associations form the essential basis for his later images and ideas, which can be symbolised by means of words.

Similarly, his manipulations of ideas, which we call thinking or reasoning, again depend on practice and experiment, and they are shaped and corrected by adult teaching and other environmental influences. His whole mental development is a cumulative process of interaction between his maturing potentialities and environmental opportunities and pressures. We can readily see, therefore, that in an abnormal or deprived environment, perceptual and conceptual development may be hindered, and consequently his intelligence level reduced; whereas in a home where rich experience of objects, words and logical thought are provided, and where emotional adjustment is harmonious, the growth of intelligence may be fostered.

Further valuable suggestions as to the brain mechanisms underlying mental development have been made by Hebb, the Canadian neurological psychologist. And he draws a most useful distinction between two uses of the term intelligence. On the one hand there is Intelligence A, that is, the basic potentiality for building up associations and for mental development generally—some property of the brain and central nervous system which is ultimately determined by the genes. Whereas the intelligence we see in everyday life or at school, and which our tests sample fairly effectively, is Intelligence B; and this represents the actual complexity and efficiency of mental processes reached through the cumulative process I have been describing. In other words, Intelligence A is the innate capacity to acquire Intelligence B. But we can never observe Intelligence A directly, nor measure it; it is merely a useful hypothesis. Thus, it would probably be wiser to forget about it and to concentrate on ' Intelligence-B-as-developed-by-environment.' Although, as shown above, the latter is very variegated, yet it can be reliably measured at any one time, and does stay fairly stable among a majority of normal children.

But the reason why it enables us to predict educability and future progress over the next few years is not so much because it is innately determined as because of its essentially

cumulative nature. By the time children reach, say, 11 years, the interaction of their innate potentialities with environmental stimulation has brought them to a certain level of intellectual functioning, and this level will largely determine their future growth, though it may always speed up, or be inhibited, in a minority depending on the kind of intellectual and emotional experiences they meet.

Intelligence and Attainment

The question will naturally be asked—how then does intelligence differ from attainments, and the answer is—much less than we commonly suppose. For innate and external factors play very much the same role in the child's acquisition of reading or arithmetic as in the building up of his intelligence. It is very natural that teachers should tend to over-estimate the contribution that schooling makes to attainment. Yet there is much experimental evidence indicating that wide differences in the school time devoted to a subject, and differences in the methods of teaching, have relatively small effects on the amount absorbed and retained. The major importance of genetic potentialities is implied by our recognition of the need for children to have reached a certain stage of maturation or reading readiness before it is profitable to embark on formal reading. Out of school experience, encouragement or discouragement at home, physical health and temperamental adjustment are also very influential, as Burt's and Schonell's investigations of backwardness have demonstrated. Thus, it is clearly false to think of school attainment as consisting simply in the absorption of information and skills that the school provides.

A further reason for regarding intelligence as a kind of attainment, and attainment as a kind of intelligence, is that there seems to be little difference in their stability or variability over periods of time. Children do, of course, vary from class to class in, say, their progress in arithmetic, but their all-round level appears to remain fairly steady in a majority, just as does the I.Q., except perhaps when they change school or take up very different school subjects.

Nevertheless, a partial distinction still seems useful. Intelligence B refers to the more general thinking capacities—comprehending, problem-solving, rational thinking, level of concept development, etc., which have very wide applicability or transfer-value to daily-life situations as well as to educational materials or to intelligence-test questions. Whereas attainments refer to more specific capacities which are more directly moulded by stimuli received in the classroom, and which probably depend to a somewhat greater extent on interest and on personality factors such as perseverance and motivation. But the distinction is a matter of degree rather than of kind, and there are many intermediate capacities which are difficult to classify. For example, vocabulary, general knowledge, and mathematical reasoning are sometimes included in intelligence, sometimes in attainments, tests.

The main corollary of this viewpoint is the abandonment of the conventional conception that intelligence is a cause of attainment, and that backwardness can be explained away if it is found that a child has a low I.Q. It is true that both Intelligence B and any

kind of attainment are largely dependent on, or caused by, Intelligence A, but as we have no means of assessing Intelligence A accurately, that does not take us any further. In the majority of children, I.Q. and average educational quotient go pretty closely together; the correlation is about 0.85. But quite a number can be found in whom one is appreciably higher or lower than the other. Surely, this should not surprise us any more than does the observation that some are better at verbal abilities, others at mathematical, or at practical, abilities. There may be certain differences in inherited talents; but apart from these we would expect different interests and experiences to produce stronger or weaker development in some abilities than in others.

The Achievement Quotient

This brings us to the topic of the Achievement Quotient, which has aroused so much controversy recently—that is the ratio of Educational to Mental Age, or Educational Quotient to I.Q. As is well-known, it is supposed to show whether children are "working up to their native capacity." If not, they are called " retarded," as distinct from merely backward, and educational psychologists or remedial teachers tend to explain the discrepancy in terms of environmental, temperamental or other factors, and to regard such pupils as the most fruitful cases for psychological or remedial educational treatment. The discussion centred around the explanation of A.Q.s above 100, i.e. children who were working " above capacity."

But there was little consideration of the points dealt with in this article—points which imply that the whole notion of working at the level expected from the I.Q. is unsound. In addition, as I have shown elsewhere, the A.Q. should be abandoned since it is a highly unsatisfactory type of measurement. It is apt to be very unreliable or untrustworthy; and inevitably it is unfair to the brighter child. Far more above than below average I.Q. children are likely to come out as retarded, although, on psychological grounds, it is the below average ones who are more likely to be failing to realise their potentialities.

How, then, should we interpret the more extreme discrepancies? When E.Q. greatly exceeds I.Q. it may be that the child is unusually docile, persevering, interested in school work, and well coached at school or at home, but is not advancing as rapidly in general understanding, nor applying his attainments to non-scholastic problems. Alternatively it may be simply that the particular kind of intelligence test does not suit him; it may be highly speeded and he prefers to work slowly, or if it is non-verbal he may lack the spatial or other ability that it involves over and above *g*.

The opposite type of discrepancy also frequently arises because the intelligence test embodies speed, spatial or other group factors in which the child is relatively strong, but which have scarcely any bearing on his educability. Nevertheless, as we know from follow-up studies of the 11+, a reliable verbal intelligence test does provide somewhat better evidence of future educability than does present attainment, when the circumstances of schooling alter. Thus, a proportion of children with E.Q.s much below I.Q.s probably does arise from environmental handicaps, emotional maladjustment, lack of interest, irregular schooling, and the like. Though these same factors are liable to

depress the I.Q. also, they may have less effect on thinking abilities which are built up to a greater extent from home and leisure experiences.

A general conclusion that follows is that intelligence tests are somewhat less valuable in educational guidance and in the diagnosis of backwardness than earlier psychologists supposed. On the other hand, the views expressed here do nothing to lessen their value in educational selection and prediction. General thinking skills as well as attainment in junior school subjects are highly relevant to grammar school success. Similarly, in streaming, or in allocating the very dull to special classes, a combination of intelligence and attainments tests is likely to be more prognostic than either separately.

Which children, then, may we regard as failing to ' work to capacity '? According to the view of this article, all children are underfunctioning to a greater or lesser degree, in the sense that neither their attainments nor their general intellectual development are as advanced as they might be, given ideal home upbringing and schooling. Social and educational policy have done much to improve this,* but the problem is still so vast that we are forced to concentrate our remedial efforts on those most likely to benefit. Now, in general, the most serious cases are those whose attainments lag furthest behind their chronological ages.

But, in addition, the skilled teacher or the educational psychologist can often discover more hope of improvement in some very backward children than in others. Unfortunately, we have to admit nowadays that an I.Q. on a verbal test, while of some relevance in picking out these cases, is less helpful than was previously supposed; and a non-verbal test is seldom of any use at all. Research evidence gives very little support for the claim that higher I.Q. children overcome their ' retardation ' more rapidly than low I.Q. children; and it has not yet revealed any more scientific means of diagnosis than the teacher's or the clinic's subjective judgment. Of course, the child with I.Q. 120, E.Q. 100 often appears to respond more readily to treatment than another of the same age with I.Q. 70, E.Q. 80, because he is much less seriously backward and easier to make contact with. It might be argued, too, that he is more worth treating as being potentially a more valuable member of the community. Yet, so far as our present knowledge goes, the latter might gain as much as, or more than, the former from careful diagnosis of his difficulties and individual attention.

One other practical point to emphasise is that non-verbal or pictorial group tests, and individual performance tests, are much inferior to verbal ones in predicting educability (except, possibly, in predicting ability for science and technical courses among older pupils). Thus, their use for scholastic purposes should be avoided. They may appear to give a fairer chance to very young children, or to older pupils who are poor readers. But with few exceptions (such as testing the deaf), they do not yield a better index of ' innate potentiality '. Group oral tests should, therefore, be used when individual Binet testing is impracticable.

*Thus, in the United States the average army recruit in World War II scored the equivalent of some 10 I.Q. points higher than in World War I. While this may partly reflect the greater familiarity of intelligence tests in the 1940's, it is probably mainly due to nation-wide improvements in the general educational level.

References for Further Reading

ANASTASI, A., *Differential Psychology* (rev. edit.). New York: Macmillan, 1958.

CLARKE, A. D. B., *et al.*, " Cognitive and Social Changes in the Feeble-minded "—Three Further Studies. *Brit. J. Psych.*, 1958, XLIX, 144-157.

HEBB, D. O., *The Organization of Behaviour*. New York: Wiley, 1949.

O'CONNOR, N. and TIZARD, J., *The Social Problem of Mental Deficiency*. London: Pergamon Press, 1956.

PIAGET, J., *The Psychology of Intelligence*. London: Routledge & Kegan Paul, 1950.

PIDGEON, D. A., YATES, A., BURT, C., *et. al.* The Relationship between Ability and Attainment. *Bull. Nat. Found. Educ. Res.*, 1956-7. Nos. 8-10.

VERNON, P. E., " The Relation of Intelligence to Educational Backwardness." *Educ. Rev.*, 1958, XI.

VERNON, P. E. (edit.), *Secondary School Selection*. London: Methuen, 1957.

YATES, A. and PIDGEON, D., *Admission to Grammar Schools*. London: Newnes, 1957.

TRANSFER AT ELEVEN PLUS—

A Summary of Evidence provided by Research

by ALFRED YATES and D. A. PIDGEON

Introduction

THE arrangements made by local education authorities for the transfer of children from primary to secondary schools are carefully and sometimes anxiously scrutinised by teachers, parents, and members of the general public. The decision to transfer a child to a particular type of secondary school not only determines the kind of curriculum that he will pursue but may also affect his chances of proceeding to some forms of higher education, and help to define his social status and the range of his vocational opportunities. For these reasons the eleven plus is generally regarded as a process of selection and popular appraisal of the methods used by local education authorities tends to be based on other than purely educational grounds. In spite of the efforts made by authorities to represent and indeed to treat transfer at eleven plus as one phase in a continuous process of educational guidance, it has become a question in which social and political considerations are involved, with the result that public discussions of the problem are often emotionally charged and there is a tendency to confuse a number of issues that in the interests of clarity should be kept distinct.

It is legitimate, for example, to ask each of the following questions: to what extent are current methods successful in attempting to discriminate between children suitable for grammar, technical and modern schools or courses as these are organised at present?; how far does the provision of schools or courses of these kinds satisfy the educational needs of the children concerned?; is it in the best interests of society as a whole to organise secondary education in this way? The first of these questions poses an essentially technical problem. The second question concerns the relationship between, on the one hand, the abilities, attainments and interests of primary school leavers and on the other, the opportunities afforded for their development by existing secondary school courses. The third embraces the more comprehensive problem of determining the needs of modern society and of examining the adequacy of our present educational arrangements to satisfy them.

In current discussions of the problems of transfer from primary to secondary schools, arguments relevant to the consideration of one of these questions are often used to answer another. The logical error that occurs most frequently consists in levelling criticisms against procedures of selection that are in fact based on objections to the ends that these procedures are designed to serve. Some parents view with disquiet the fact that what appears to them to be a largely irreversible decision affecting their children's educational and vocational prospects is made when the children are ten years of age. The target for their criticism, however, is often the examination—or some component of it—on which the decision is based.

We have provided this introductory comment on the circumstances in which judgment tends to be formed of the efficiency of the arrangements for transfer at eleven plus

in order to define the scope of this article. The research evidence available is almost wholly concerned with the first of the three questions listed above. In the nature of things it is existing arrangements rather than proposed reforms that can be subjected to rigorously controlled inquiry. The evidence that we propose to summarise, therefore, is derived from researches that have examined the operation of arrangements for allocating children to secondary school courses within what has been termed the tripartite system. Since there have been very few researches of a large-scale or intensive kind into selection for technical courses the bulk of the evidence is, therefore, related to the problem of how most effectively to discriminate between those children who appear to be suitable for a grammar school course and those who do not. It is only to this narrow and specific question that an answer based on a wealth of research evidence can be provided. The broader questions concerning the kinds of secondary school organisation that could best serve present-day needs have not yet been systematically examined.

In the summary that follows we have not, except in a few instances, referred to the specific researches that have contributed to the evidence, nor entered into any discussion of the design and conduct of the inquiries or the methods of analysis used. For those who wish to consult the evidence at first hand we have provided a bibliography containing a list of the sources consulted.

The contents of the eleven plus examination

We propose to use the term ' eleven plus examination ', in accordance with popular usage, as an umbrella term to cover all the procedures—whether these consist in fact of written examinations and tests or of interviews, teachers' assessments or other devices—which education authorities employ as a basis for their allocation of pupils to different types of secondary school course.

Since each authority is free to devise its own form of examination and there is, consequently, a large variety of procedures in use, it is imprecise to speak of ' the ' eleven plus examination. This variety, however, consists for the most part in different combinations of a relatively small number of possible components and in different ways of treating the results that they yield, and so by considering the evidence relating to each of the parts separately and to the more commonly employed combinations, a reasonable estimate can be formed of the efficiency of the procedures as a whole.

1. *Intelligence tests*

Pride of place must go to the intelligence test both because this is the feature of the eleven plus examination which has occasioned a good deal of controversy and because its use, in spite of the criticisms levelled against it, is widespread. (At the time of the last survey carried out by the Foundation in 1956 it was found that just over 88 per cent. of local education authorities were using intelligence tests as part of their eleven plus procedures.) Numerous researches have shown that intelligence tests predict success in grammar schools more closely than any other type of written test or examination that has as yet been tried.

In spite of the fact that these results have been widely publicised there is still a considerable amount of opposition to the use of such tests. Part of this undoubtedly stems

from the mistaken notion that they yield, in some mystical fashion, a measure of innate ability. It is far less wounding to his pride for a parent to be informed that his child's attainments in arithmetic are below a required standard—this can be attributed to misfortune or bad teaching—than to be told that his child has failed to satisfy the examiners in a test which is popularly supposed to reflect his degree of inherited ability. Neither of the two organisations—the National Foundation and Moray House—that supply the bulk of the intelligence tests set by local education authorities uses the word intelligence in the title of its tests. Moray House has recently changed the description to ' Verbal Reasoning '. The Foundation, from the outset, avoided any trace of the tendentious by styling them simply as ' Verbal ' tests. It is unlikely, however, that the popular use of the term ' intelligence tests ' as a description of these instruments will disappear for some time and the associated idea that they involve the measurement of some basic and largely inherited quality will also probably linger.

A somewhat more serious objection to the use of intelligence tests in the eleven plus examination is that research has shown that they are susceptible to the effects of coaching and practice. As a consequence many children are involved, in or out of school, in preparatory activities which take up time that could perhaps be more profitably spent in other pursuits. This objection, probably as a result of the prejudice against the supposed nature of intelligence tests has been considerably exaggerated. Indeed, this charge is sometimes brought against intelligence tests without acknowledgement of the proven fact that all the other kinds of test or examination that are or can be used in the eleven plus examination are equally affected by special preparation.

As far as the predictive value of intelligence tests is concerned, coaching is only likely seriously to distort the outcome if some candidates have been expertly prepared for the tests whilst others are totally unfamiliar with them. This situation is hardly likely to arise nowadays. The fact that the effects of coaching are minimal unless the material used for the purpose is similar in content and format to that included in the test proper has been amply demonstrated. Moreover, it has also been shown that the advantages accruing to those who have been expertly coached can be offset by arranging for all children who are about to enter for the examination to be allowed opportunities to familiarise themselves with the type of test to be encountered.

The only valid objection remaining against the use of intelligence tests, therefore, is the fact that if they are used in the authority's examination they are likely to be introduced into the curricula of some primary schools. This is tantamount to saying that they have what has come to be called a ' backwash ' effect on the work of the primary schools. Again this is a criticism that can be levelled against other tests or examinations. Whatever form of examination is used at the end of an educational course, it will tend to influence the content of the curriculum and the methods of teaching employed.

In deciding whether or not intelligence tests should be used in an eleven plus examination, therefore, the major consideration should be whether or not the particular kind of backwash effect associated with these tests is sufficiently serious to warrant the exclusion from the examination of one of its most efficient components. We believe that some, at least, of the activities in which primary school children are involved when they are being coached for an intelligence test—discovering synonyms, opposites and analogies,

doing simple exercises in verbal reasoning, breaking codes, analysing number series, etc.—would find a place in the curriculum, even if intelligence tests were no longer included in the eleven plus examination.

2. *Tests of attainment*

The volume of popular criticism levelled against the use of intelligence tests has tended to obscure the fact that tests of attainment in English and arithmetic are used in most examinations, and since equal weight is usually given to each of the three parts of the examination, a child's level of attainment is a more decisive factor than his performance in the intelligence test. The majority of authorities employ standardised objective tests of attainment, but there has been a tendency in recent years for the reintroduction of the older forms of examination, in English particularly, sometimes in place of standardised tests but more often in addition to them. The reasons for this apparent ambivalence—the wish to restore the older form of examination and yet to retain the newer type of objective test—are first, that considerable dissatisfaction has been voiced, particularly by those who teach in grammar schools, about the effects on the attainments of primary school children of the exclusive use of objective tests; and, secondly, that objective tests have been demonstrated to be both more reliable[1] and more valid[2] than their predecessors.

The arguments against the use of objective tests are familiar enough. They refer to the ' backwash ' effect of tests which rely in the main on ' multiple choice ' items—questions for which a series of possible answers are supplied, the examinee being required to indicate the one acceptably correct response. It is often alleged that if such tests are used to assess attainment in English, children in primary schools tend to be deprived of adequate opportunities to practise writing in consecutive prose. Although evidence to support this charge is somewhat meagre it has, nevertheless, achieved a widespread currency.

There is considerable evidence relating to the relative efficiency of objective tests and the older forms of examination. A follow-up study* carried out in the area of an authority which employs both types of examination showed that the prediction afforded by the objective tests alone was as high as that for the total examination. This result can be justifiably interpreted in two ways. Since objective tests alone achieved as good a result as a procedure involving the addition of older forms of examination, the latter could be dispensed with, thus reducing the amount of time and money spent on the process. Alternatively, since the addition of the older forms of examination did not *reduce* the efficiency of the examination, they may safely be included in order to offset the supposedly adverse ' backwash ' effects of the objective tests.

An attempt has been made to devise a new kind of English test that makes the best of both worlds, so to speak. Such tests, although constructed on the pattern of modern objective tests in that they contain considerably more items than the traditional type of examination, nevertheless follow the example of the latter in requiring creative responses,

[1]The reliability of a test is the degree to which it is likely to yield consistent results on successive occasions.
[2]The validity of a test is the extent to which it achieves its purpose—i.e. in this instance, affords a prediction of success in a grammar school.
*See Yates and Pidgeon (1957).

some including the construction of complete sentences. It may be supposed, therefore, that the ' backwash ' effect of these new type tests is not dissimilar to that of the older examinations, yet they have been shown to be as reliable as the purely objective tests and to enjoy even a slightly higher level of predictive efficiency.*

English essays have also been reintroduced into the procedures used by many authorities in recent years. The reason for doing so is, no doubt, to encourage those primary school teachers who, in preparing their pupils to face objective tests, are alleged to follow an unduly restrictive curriculum. The researches of Wiseman (1949) and others have encouraged authorities to take this step, by demonstrating that it is possible to employ methods of marking essays which yield satisfactorily reliable results.

3. *Teachers' assessments*

It can scarcely ever have been doubted that the teachers who have had charge of a child's primary education are able to make a sound appraisal of his level of ability and attainment and can, therefore, predict with some accuracy his chances of success in a grammar school course. The major problem has been that of discovering a satisfactory means of collating the assessments provided by teachers from different schools. It has been demonstrated that teachers can satisfactorily rank their own pupils with respect to their suitability for a grammar school course.

It is clearly important, however, in awarding grammar school places to children from different primary schools, to make due allowance for the fact that the proportion of suitable candidates in one school is not necessarily the same as that in another. Various devices, designed to make this necessary adjustment, have been tried. The quota scheme, pioneered by the Walsall authority, the methods of scaling and quantifying teachers' assessments introduced by Sutcliffe and Canham (1945) and by the National Foundation (Yates and Pidgeon, 1957), and the procedure devised by G. F. Peaker and tried out in part of the West Riding, represent attempts to place on a comparable basis assessments made by teachers in different schools. All but the last of these makes an adjustment based on test scores or examination marks. The West Riding procedure relies on the observed stability of the proportion of successful candidates in each of the primary schools concerned.

There is a considerable body of evidence to show that if teachers' assessments are suitably adjusted in one of the ways described above they yield a predictive measure equal to that provided by an examination and, indeed, superior to that provided by any single test. This fact has raised the important question as to whether it might be possible to dispense with an eleven plus examination and to rely wholly on the assessments of their pupils provided by primary school teachers. The available evidence suggests that the combination of teachers' assessments and a soundly constructed examination is more efficient than either used alone. It may well be possible, however, by improving the methods whereby teachers' assessments are made and utilised to reach a level of predictive efficiency that could only be improved marginally, if at all, by the addition of an examination.

*Op. cit.

Some take the view that the use of teachers' assessments, without an examination, would be desirable even if there were some consequent loss of efficiency (the evidence suggests that this would be slight) on the grounds that primary schools would be liberated from the restrictive influences of the examination. Against this it is necessary to set two considerations: in some areas, particularly those in which competition for grammar school places is severe, teachers may not be willing to shoulder the whole burden of allocating pupils to different types of school; the public, generally, has a higher regard for procedures of selection that are based on apparently impartial examinations than for those which are more directly dependent on human judgment.

The problem of the borderzone

No known form of examination or assessment has achieved perfect reliability. The practical consequence of this is that if any eleven plus examination were to be repeated—the tests given on a different day, for example, or the teachers' assessments requested in a different form—there would be some inevitable variation in the results. In other words, drawing a line across a list, on whatever basis the list has been compiled, in order to distinguish between those to be admitted to a grammar school and those to be excluded cannot be justified. For this reason it is generally recognised that the dividing line should be replaced by two lines, dividing the candidates into three groups: those about whose suitability for a grammar school course there can be no reasonable doubt; those who are clearly unsuitable; and that group of candidates, usually described as belonging to the borderzone, about whom no firm prediction can be made on the basis of the evidence available.

The problem of delimiting the border-zone is one that can be satisfactorily solved by discovering empirically either the degree to which the tests and assessments used fall short of perfect reliability, or the accuracy with which they predict future performance. The problem of what further assessments to make of the potentialities of the candidates within it in order to arrive at a more satisfactory prediction is far less tractable. The law of diminishing returns stubbornly frustrates attempts to achieve the fine discriminations that are required at this stage in the selection procedure. One self-evident principle, which is nevertheless frequently ignored, is that in dealing with border-zone candidates techniques of assessment should be sought which are more comprehensive and effective than those which have achieved the initial relatively coarse groupings. It is scarcely reasonable to expect, for example, that a brief interview can enable discriminations to be made which a combination of examination scores and assessments provided by teachers and based on continuous and close observation over a period of years have failed to achieve.

The evidence suggests that to improve on the prediction afforded by a sound initial procedure—based, for example, on suitably scaled teachers' assessments and a battery of standardised tests—requires the collection of a considerable amount of additional evidence concerning the child's abilities and attainments and a procedure whereby this evidence can be reviewed globally by an experienced panel of judges. Even an elaborate and time-consuming procedure of this kind has been found to yield only a small improvement.

The effects of differences in age

There is evidence to show that differences in age, even those amounting to no more than a few months, affect the performance of children in examinations and tests of all kinds. Thus, if no allowance were made for these differences, there would be a tendency for the older children within the age group considered for admission to grammar schools, to gain a disproportionate share of the available places. The use of standardised, objective tests, counteracts this tendency in that appropriate allowances for differences in age are incorporated in the system of transforming the ' raw ' or actual marks gained by a child into the standardised scores which are used to express his relative level of performance.

Other forms of assessment, however, are not necessarily free from these effects. Many of the authorities that use unstandardised examinations make some adjustments to the scores to offset the injustice that would otherwise be done to the younger candidates. These adjustments, however, are not adequate in every instance and, moreover, it is rare to find that age allowances are considered in the case of English essays, primary heads' assessments, or interviews. Since primary heads' assessments have been shown to be of considerable value as predictors of academic success and are for this reason becoming more widely employed it is important that their main defect—the demonstrable tendency for these assessments to be biassed in favour of older children—should be remedied. A number of methods have been devised for this purpose and are described in detail in Yates and Pidgeon (1957) and Vernon (1957).

Differences in test scores between boys and girls

The task of authorities in attempting to administer an eleven plus examination that is manifestly fair to all candidates is further complicated by the fact, established by evidence from a number of sources, that the performance of girls is superior to that of boys in most if not all of the tests and examinations commonly employed. This superiority of girls appears to be maintained for some years after the children have been transferred to secondary schools. It is reasonable to suppose that it is associated with the fact that girls mature at an earlier age than boys and that therefore the differences manifest at eleven may well disappear subsequently, although there is, as yet, insufficient evidence to allow this conclusion to be drawn with certainty. In the circumstances the most satisfactory course for an authority to adopt would seem to be that of treating boys and girls separately for the purposes of allocation to secondary school courses.

The validity of the examination

The most crucial question to ask about the eleven plus examination is to what extent it serves its major purpose of distinguishing between those who are and those who are not suitable for a grammar school course. It is impossible, in the nature of things, to answer this question directly and empirically. To do so it would be necessary not only to discover a generally accepted criterion of success in a grammar school course but also to arrange for all children to follow such a course. Only by this means would it be possible to obtain direct evidence concerning the suitability or otherwise of those candidates excluded by the selection procedure. Another complication, although one that can be to some extent resolved, arises from the fact that there are marked differences between

schools of the same type, so that it may well be that a child who does not satisfy the requirements of one grammar school, for example, may well have prospered had he been allocated to another.

Three main criteria have been employed by those who have attempted to evaluate the effectiveness of the eleven plus examination: G.C.E. or School Certificate results; internal school examination marks; and the assessments provided by the secondary school teachers concerned. It may well be argued that any one of these criteria, or any combination of them constitutes an inadequate measure of the extent to which children have been appropriately allocated. On the other hand, an eleven plus examination which failed to discriminate between those children on the one hand who are able to put up a satisfactory performance in school examinations and to earn the commendation of those who teach and those, on the other, who are adjudged incapable of reaching the standards required of them could scarcely be defended. Criteria depending on examination marks and teachers' assessments of their pupils' suitability for a grammar school course may not be wholly adequate for the purpose but there do not appear to be any practicable alternatives.

The researches carried out by the Foundation used the two latter criteria, and by comparing them, for each secondary school separately, with the results of the eleven plus examination, it was possible to assess the degree to which the latter had afforded an adequate prediction. The analyses were carried out for the schools separately to allow for the fact that there are marked differences between schools even of the same type.

It is clear that if all the comparisons showed perfect agreement—that is, if the eleven plus examination placed children in exactly the same order of merit as that which was subsequently found to obtain in the secondary school to which they were transferred, it would be reasonable to assume that the children ' above the line ' on the original examination list were, in fact, more worthy of admission to grammar schools than those who fell beneath it. On the other hand, it also follows that if a lack of perfect correspondence is discovered then there must have been some error in the initial allocation, and the extent of this error can be estimated by observing the degree of correlation between the eleven plus results and the subsequent performance of the children concerned.

The Foundation's estimate of the size of this error—an estimate which is in agreement with that made by other workers [see Vernon (1957)]—was that, with the best available procedure, some ten per cent. of the candidates are likely to be ' wrongly ' allocated. ' Wrongly ' is placed in inverted commas because this statement does not necessarily imply that, as a result of this apparent error, the children concerned are likely to receive educational treatment inappropriate to their needs. The differences between children who form the border-zone are often very small. The child who is allocated to a secondary modern school, but who is nevertheless slightly superior academically to one who just manages to secure admission to a grammar school, is not necessarily deprived of opportunities to develop his potentialities. Nor is it by any means certain that his slightly weaker colleague will derive no benefit from his experience in a grammar school.

On the other hand, since errors of this order are likely to be made by even the most efficient procedures that have as yet been devised, the policy, now adopted by more than

half the education authorities, of so modifying the organisation of secondary education as to blur the hitherto sharp distinctions between grammar schools and other types of secondary school would seem to have much to commend it.

There is little doubt that the eleven plus examination, as developed by many authorities, is the most efficient examination to be found in our educational system. It is almost certainly more reliable than any of the examinations that children encounter at later stages in their careers and, as a process of selection, it is probably more valid than those employed for the selection of entrants to Universities, the Civil Service, or the various professions. The main criticism that can justifiably be brought against it is that it has at times been required to accomplish the virtually impossible task of segregating into courses which are sharply differentiated, children between whose abilities, attainments and ascertainable educational needs there is a scarcely perceptible difference.

Select bibliography

BOSOMWORTH, G., " The Use of Teachers' Assessments in Allocating Children to Secondary Schools." *Durham Res. Rev.*, Vol. IV, 1953.

BURT, C., " Validating Tests for Personnel Selection." *Brit. J. Psych.*, Vol. XXXIV, 1943.

CLARK, H., " The Effect of a Candidate's Age upon Teachers' Estimates and Upon His Chances of Gaining a Grammar School Place." *Brit. J. Ed. Psych.*, Vol. XXVI, 1956.

DEMPSTER, J. J. B., *Selection for Secondary Education*. London: Methuen, 1955.

EMMETT, W. G., *An Inquiry into the Prediction of Secondary School Success*. London: University of London Press, 1942.

EMMETT, W. G., and WILMUT, F. S., " The Prediction of School Certificate Performance in Specific Subjects." *Brit. J. Ed. Psych.*, Vol. XXII, 1952.

EMMETT, W. G., " Secondary Modern and Grammar School Performance Predicted by Tests Given in Primary School." *Brit. J. Ed. Psych.*, Vol. XXIV, 1954.

FLOUD, J. E., *Social Class and Educational Opportunity*. London: Heinemann, 1957.

FREEMAN, F. S., and MILES, C. C., " Sex Differences." *Encyclopedia of Educational Research*, Rev. Ed. London: Macmillan, 1950.

HARTOG, P. J., and RHODES, E. C., *An Examination of Examinations*. London: Macmillan, 1935.

HARTOG, P., RHODES, E. C., and BURT, C., *The Marks of Examiners*. London: Macmillan, 1936.

JAMES, W. S., *et al.*, *Studies in Selection Techniques for Admission to Grammar Schools*. Bristol Univ. Inst. Educ. Publ. No. 3. London: Univ. of London Press, 1952.

MCCLELLAND, W., *Selection for Secondary Education*, 1st Imp. London: Univ. of London Press, 1942.

MCINTOSH, D. M., *Promotion from Primary to Secondary Education*. London: Univ. of London Press, 1948.

MCINTOSH, D. W., WALKER, D. A., and MACKAY, D., *The Scaling of Teachers' Marks and Estimates*. Edinburgh: Oliver & Boyd, 1949.

MINISTRY OF EDUCATION, *Early Leaving: A Report of the Advisory Council for Education (England)*. London: H.M.S.O., 1954.

NATIONAL UNION OF TEACHERS, *Transfer from Primary to Secondary School*. London: Evans Bros., 1949.

PAULEY, F. R., " Sex Differences and Legal School Entrance Age." *J. Ed. Res.*, Vol. XLV, No. 1, 1951.

PEEL, E. A., and RUTTER, D., " The Predictive Value of the Entrance Examination as Judged by the School Certificate Examination." *Brit. J. Ed. Psych.*, Vol. XXI, 1951.

PILLINER, A. E. G., " The Position and Size of the Border-Line Group in an Examination." *Brit. J. Ed. Psych.*, Vol. XX, 1950.

RICHARDSON, S. C., " Some Evidence Relating to the Validity of Selection for Grammar Schools." *Brit. J. Ed. Psych.*, Vol. XXVI, 1956.

RUTTER, D., " An Inquiry into the Predictive Value of the Grammar School Entrance Examinations." *Durham Res. Rev.*, Vol. I, 1950.

SUTCLIFFE, A., and CANHAM, J. W., *Selection for Secondary Education*. London: John Murray, 1945.

Times Educational Supplement, " Selection at Eleven—III. Tyranny of Tests. Yorkshire System of Exemption." 23rd September, 1955.

VALENTINE, C. W., and EMMETT, W. G., *The Reliability of Examinations*. London: Univ. of London Press, 1932.

VERNON, P. E. (Ed.), *Secondary School Selection.* London: Methuen, 1957.
WALKER, A. S., *Pupils' School Records.* London: Newnes, 1955.
WATTS, A. F., and SLATER, P., *The Allocation of Primary School Leavers to Courses of Secondary Education.* First Interim Report. London: Newnes, 1950.
WATTS, A. F., PIDGEON, D. A., and YATES, A., *Secondary School Entrance Examinations.* Second Interim Report. London: Newnes, 1952.
WIGFORSS, FRITS, *The Entrance Examination in View of Later School Performances.* Stockholm: P. A. Norstedt & Soners Förlag, 1937.
WISEMAN, S., "The Marking of English Compositions in Grammar School Selection." *Brit. J. Ed. Psych.*, Vol. XIX, Part III, 1949.
WISEMAN, S., *Problems Connected with the Selection of Children for Secondary Schools.* Ph.D. Thesis, University of Manchester, 1952.
WRIGLEY, J., "The Relative Efficiency of Intelligence and Attainment Tests as Predictors of Success in Grammar Schools." *Brit. J. Ed. Psych.*, Vol. XXV, 1955.
YATES, A., and PIDGEON, D. A., *Admission to Grammar Schools.* Third Interim Report. London: Newnes, 1957.

THE WISH TO LEARN

Research into Motivation

BY W. D. WALL

I. Introductory

The Will to Learn

THE teacher is directly and continuously concerned with learning and with the will to learn. He frequently observes that a child who finds it genuinely difficult to remember a minimum of facts necessary, say, to the comprehension of geography will commit to memory without effort equally difficult facts concerning aircraft or ' pop ' singers. He probably notices, too, that inside and outside the classroom, his pupils acquire a good deal of knowledge, a great many techniques, and attitudes which are only incidentally related to what they have been specifically taught at home and at school or even to what they may themselves have deliberately set out to acquire.

Educationally, we are faced with a major problem. What is this will to learn and how do we direct it to those bodies of knowledge, those techniques and attitudes which it is the business of the teacher to impart? How can we see that it leads to incidental learning of the right kind? What is it in the original learning situation, or subsequently, that leads to retention of what has been learned? What is it in one piece of learning that facilitates or hinders the pupil in his subsequent learning?

Primary Motivations.

All this directly concerns what the psychologist calls ' motivation '—perhaps the most obscure and difficult of all the theoretical and practical issues in general and educational psychology. It does, however, seem to be very clear that the concept of motivation covers a variety of aspects of human (and animal) behaviour and that, for clarity of thinking and for practical use, it is as well to make a few distinctions at the outset.

Some motivations are evidently primary—that is to say, that they spring from more or less crude biological drives like hunger, thirst, sex and so on. In their primitive forms, they certainly directly and powerfully motivate human behaviour and the more they are denied, the greater becomes the tension in the organism until they are satisfied. It is certain, too, that a drive like hunger will lead an animal or a human being to learn very rapidly the ways by which it can be appeased and that such learning is likely to be durable.

The primary drives of hunger and thirst seem to be the main springs of learning in the first few days and weeks of life; but even they rapidly become, under normal circumstances, somewhat sophisticated and overlaid with other learnings and developments of ' secondary' or ' derived ' motivations. More and more, as the child grows older, these secondary motivations become important and it is difficult to see a primary drive as the mainspring of any learning which he may do in school. Indeed, we are often tempted to see such primary motivations as hindering factors. Hunger, fear, aggressiveness, the search for security and love, if they reach any level of intensity in a child of school age may well

lead to behaviour and learnings which run counter to what the teacher is trying to teach or what the parents wish from their child. So far as, for example, learning to read or to calculate are concerned, the appeal is to some such motives as interest, self-assertion through success, affection for the teacher and the like which are derived or secondary, that is, they are themselves learned and elaborated out of previous learning experience.

Unconscious Motives

We have so far spoken as if the will to learn were always conscious. This, however, is far from being the case, even with the intelligent and sophisticated adult. It becomes painfully apparent when we try to teach a child who has, for example, developed an emotional difficulty and fear of arithmetic. He may consciously wish to succeed; and he may try and try hard: but he evidently has an unconscious counter motivation which defeats his efforts and ours. Similarly, there is little doubt that children and adolescents learn rapidly and apparently without conscious effort many difficult things which satisfy needs which they could not formulate consciously. In this form of learning we may surmise, from work on child development and from the studies of maladjusted children, that such basic drives as the search for affection and acceptance, the need for security, the desire to assert oneself and the sex-drive, particularly in adolescence, are of primary importance.

Incentives

Clearly, there are some circumstances in which these inner motivations, whether they are conscious or not, are sufficient in themselves to lead to action, to seeking behaviour the end of which is to satisfy the need; and clearly, if the environment prohibits certain easy ways of satisfying the need, the child (or the animal) will learn the relatively more difficult techniques fairly rapidly.

It is on this basis that much of the early learning and training of infants takes place. The infant seeks to satisfy its hunger; the mother provides the satisfaction but at the same time deflects the child from one form of behaviour to another. Similarly, the infant seeks the warmth and physical security which nursing provides and mother tends to use this as a reward for steps forward in behaviour and self-control which the child learns. Thus, from a very early age, parents, and particularly mothers, tend to use *incentives*: that is, the infant is stimulated and motivation is aroused by something extrinsic to himself. This, however, as every mother quickly finds out, will only be successful if the incentive used really does stimulate some need in the child and if there is no competing need of greater strength.

In the class-room, this interaction between more or less primary motivations and the incentives which the teacher offers, is even more complex than in the early years of life and the interaction is less evidently direct. We may illustrate this by the frequently recurring situation in which the teacher offers as an incentive praise which is to be earned for work well and successfully achieved. Such praise means, to the five-year-old child, say, affection and the security of being loved. Thus, the anticipation of it is a stimulus to his search for acceptance and security. Thus far the situation is likely to arouse effort to do what the teacher wishes him to do.

But there may be a competing need. The child may be afraid of the teacher; or he may have sat still already for so long and the need for physical movement is greater than anything else. The incentive of praise will not work if other motives are operating more powerfully. Nor will it work if the task which has to be undertaken is beyond the child's capacity. As we shall see later, if strong motivation is aroused in the face of an impossible task, behaviour may become more disorganised in proportion to the strength of the individual drive. Indeed, we may say that conflicting drives and a disproportion between the motivation aroused and the possibility of achievement are the two circumstances which do most to disorganise learning in children and which are most likely to lead, if they are repeated, to serious difficulties in growth.

Group Influences

In very early life, it seems likely that all motives are individual and personal and probably in the early school years this may continue to be largely true. However, children very rapidly come to form a part of the class group and of sub-groups within the class or within the school. All teachers sooner or later become aware of the fact that there is such a thing as group motivation. The stimulus which the teacher provides to learn may well operate differently according as to whether it is given in a direct individual context or whether it is given to the class as a whole. Not infrequently, to our regret, we find that the class or a section of it produces a powerful counter-motivation to what we are trying to do; and there is a kind of solidarity which no pupil seems willing to break.

Conversely, if once a group accepts and makes its own a motive put forward by the teacher, the learning which is undertaken will be strikingly effective even for those members of the group who, individually, would find the work beyond their powers. It seems probable that, in the junior school and onward until well into adolescence, this social or group motivation, as we may call it, is more important educationally with normal children than is anything more evidently individual and direct.

Clearly, motivation in human learning is a highly complex thing, difficult to study objectively, and instinct theory, in its earlier form, even when elaborated by secondary or derived motivations, does not suffice to account for everything, even in the learning of children. There is a further important notion, first formulated by Allport, which we must add and it may be that it is, educationally, the most significant of all. While it is true that much human behaviour can be accounted for as the result of biological drives of an innate, more or less instinctive kind, and of their elaborations and derivatives acquired through the process of living and learning, it is also true, as Allport points out, that some motives may be directly acquired and achieve an autonomy of their own. An obvious example is smoking. Once the habit is acquired, it will lead to all kinds of behaviour and may have much the same motive force behind it as hunger. It is a part at least of the task of education at home and at school to create desirable motives and habits which become ' functionally autonomous '.

II. Research Work

The purpose of this somewhat lengthy introduction is to make clear some at least of the complexities of the subject under study and to suggest a part at least of the reason why

much of the research evidence is difficult to interpret and to apply in the day-to-day teaching situation. In spite of the apparent paucity of direct classroom studies, however, there is much to be gleaned from the literature of research both in general and in educational psychology.

Animal Learning

Since the thirties, and particularly since the early work of Thorndike, there has been a great deal of work on the learning of animals and a consequent exploration of the effectiveness of what we have called earlier, primary motivations, like hunger, thirst, sex and the like. The general upshot of this work is, as we might expect, a confirmation that such primary motivations are extremely powerful.

It has also become clear that they are by no means the sole motivations seen in animal learning nor are they necessarily effective in the same way at different levels of intensity. It is apparent that too great an intensity is likely to lead to disorganisation of behaviour which in some cases looks like panic.

It has, too, become clear that motivation is a blend of two things: a state of inner tension which may be high or low; and the external stimulus which may be strong or weak. Motivation depends upon the relationship between the two at any given moment. It is held, too, that, in general, animal learning is the result of conditioning: that is to say, that a successful response is learned because it reduces the need or motivation; and an unsuccessful response is eliminated because it does not lead to the satisfaction of a need. However, while this is generally true, certain experiments have shown that ' punishment ' (i.e. usually electric shock in these experiments) can and does lead to learning.

Curiosity and Manipulation

Crudely, these findings concerning primary motivations apply to human learning, especially in the learning of babies; but they do not suffice to account for learning in the school situation or indeed for most of what occurs after the first few months of life, particularly in the intellectual field. There are two other findings from the work with animals which are more directly relevant. Some at least of the studies of ape learning have demonstrated that there are motivations other than the primary drives or needs. When these primary drives are at a low level, i.e. when the animal is satisfied and not fatigued, something which might be described as ' curiosity ' occurs and the animal will explore more or less at random. Similarly, too, apes will show ' pleasure in manipulation '—they will repeat successful activity with puzzle apparatus not now for a direct reward but apparently for the sheer pleasure of doing it.

Observations of pre-school children tend to suggest that ' curiosity ' and ' pleasure in manipulation ' are two of the most important factors in learning and in the mastery of skills or knowledge, which may have been acquired under other stimuli—and this even from the earliest months of life. Babies tend to play with a skill once acquired. Finally, it has been shown in a number of experiments with animals that there is such a thing as incidental learning—that is, learning which is apparently unmotivated—though it is difficult to demonstrate absence of motivation in animals and *a fortiori* in humans.

Insight and Feed Back

In most of the conditioning or learning experiments undertaken with animals the process exhibited by the animal is the laborious one of 'trial and error'. The rat running in the maze, the monkey in a puzzle cage, continue more or less at random to try to get at the bait, slowly eliminating errors, but often without that leap forward which comes when insight occurs and enables it to short cut the trial process. Experiments with human maze learning and some experiments with animals have shown, however, that insight does occur and that some learning is not simply conditioning in the old trial and error sense. Evidently in the teaching situation 'insight'—which may be the pupil's own more or less unaided discovery of a short cut to the solution of a problem or the result of skilful teaching—is of primary importance. In the context of the study of motivation it is equally so since the taste of sudden success and the insight itself tend not only to reinforce the original drive but may provide a notice on their own account and stimulate further effort.

Animal psychology can throw no light on certain further aspects of our topic which are of fundamental importance. Unconscious motivations, the relationships which may well exist between motivation and incentive on the one hand and the subject's level of aspiration—that is to say, the goal which he sets himself—on the other, and the generally greater complexity and indirectness of human functioning, make direct inferences from animal studies to human psychology hazardous.

Basic Needs of the Personality

Certain things are, however, clear. The concept of basic drives and of 'need reduction' is of the greatest utility when we are considering infants and children of pre-school age; and certainly much if not all early learning is motivated in this way. However, from the outset of life sheer manipulation and 'curiosity' are manifested by babies whose other needs are satisfied and the early stages of play as a means of learning and expression and of the construction of symbolic thought are assigned by Piaget to this.

From the point of view of the child of school age, however, the principal fact of importance is that curiosity, interest and the desire to manipulate and explore can hardly manifest themselves unless the basic needs of children are more or less continuously satisfied. It goes without saying that such things as hunger, fatigue, and marked insecurity are liable to inhibit forms of activity which are not directed to their removal or satisfaction; and we know from several generations of work with maladjusted children that a denial of the need for affection, of acceptance, of the need for an outlet for aggression will sooner or later lead to a failure in school learning. It is frequently not so clearly seen that such basic needs of the personality are for normally developing children an important part of the whole context of school learning, often in ways which are difficult to analyse and which differ somewhat from age to age.

When the child begins school it seems likely that he views his teacher in much the same way and with something of the same attitudes as he does his mother or father. The praise and blame of the infant mistress, her acceptance or rejection, are important motives which will in large part determine his behaviour and his eagerness to learn. It seems probable, from the numerous studies which have been made of teacher characteris-

tics and of the relation between liking for a teacher and liking for the subject which he teaches that such direct personal interaction of pupil and teacher remains an important source of motivation at least until well on into adolescence.

Nevertheless, the normal progression of growth should be that, because the basic needs of the personality are satisfied in the child's home life and generally in school, his energy is steadily freed to become occupied with the objective work. This ' liberation ' of the self from the self ', to quote Hamley's phrase, can only proceed steadily provided that outlets, of a legitimate and constructive kind, are offered for aggression and self-assertion and provided, too, that home and school environment generally favour increasingly social and heterocentric attitudes. In other words, school learning depends mainly upon secondary and derived motivations which can only be powerful stimuli of use to the teacher if the pupils are satisfied in other more basic directions.

In this, certain broad rhythms of development must be taken into account. In early school years primary motivations—love, aggression, fear—may tend to predominate in many children, though side by side with them we can see evidence of much curiosity and activity for its own sake. Normally, in the junior school period, the emphasis seems to shift to secondary motivations; children enjoy acquiring skills and knowledge, particularly if their desire for mastery and achievement is fostered by praise and success. Later, the onset of puberty and adolescence seems, for many, to sharpen the demands of primary needs, particularly self-assertion, sex and the desire for security, this time of an emotional—social kind.

It is often difficult for the teacher of young adolescents to stimulate them to learning of the conventional type; their interests and their needs lie elsewhere. It is questionable whether it is wise to persist in somewhat arid reliance in abstract curiosity, disinterested desire for achievement and the like in such cases. Probably a recognition of the preoccupations of the growing boy and girl, of their personal and social problems, of their fears and anxieties (many of them homely enough and clumsily formulated if at all) and the adaptation of curricula and method to them would produce more effective and rapid learning.

In the second half of the ' teens ' usually motivation appears to change. By this time many boys and girls have left school and the realities of life appear to provide acceptable stimuli to effort. Not infrequently, the ' lazy ' and ' apathetic ' youth of 15 turns into the young man or woman who is working hard at evening school or technical college or catching up his apparently lost years.

It is against this broad background that we should consider and interpret the evidence which we have as to the effectiveness of the various ' secondary ' or ' acquired ' motivations.

Curiosity and Interest

It will be recalled that the work with animals has shown that purely manipulative and exploratory activity and apparent curiosity occur even in rats and more so in monkeys. It seems, too, that the amount of this as well as its apparent force as a motive, increases as one goes up the scale in species and within species according to intelligence. Experi-

enced teachers are well aware that the more intelligent children are likely to display more interest and more curiosity than their less able contemporaries. Equally, such intelligent children are likely to have a wider and more mature range of interests. The problem is to know the precise psychological function of curiosity and interest and the means of rousing and sustaining them as motives for educational purposes.

There are a great many studies of the interests of children and adolescents at various ages which tend to show the existence of general tendencies in the development of interest, some sex differences in interests especially in the ' teens ', and a wide range of individual variation. When the declared interests of a given child or group of children are closely examined one tends to find that the overt content corresponds to some deeper and more or less permanent personal need which may find expression in many different ways at different times. Similarly, the same expressed interest may mean very different things in different children of the same sex and age. Hence, the advice often given to teachers to make use of the interests of pupils as a motivation to school learning needs some qualification and interpretation.

The work of Gardner, Daniels and others in England, and of the Ecole Nouvelle (Freinet, Cousinet and others) abroad, though little of it is in the nature of rigorously controlled experiment, nevertheless shows how effective the expressed interests of children may be to secure initial motivation for a task. Such work also, however, shows that the child's interest and curiosity alone cannot be relied on to carry him through the learning process without other stimulation and encouragement. It would seem that the teacher's task is to accept the declared interests and curiosity of his pupils as a starting point and indeed to use them as far as possible all along the way. At the same time, however, she must so arrange the environment that one interest provokes another and that the feeding of interest or the satisfaction of curiosity are made sufficiently easy so that the child can with reasonable effort experience success, fairly continuously.

That this is so is abundantly clear from the detailed information given by Susan Isaacs in her *Intellectual Development of Young Children.* The environment of the Malting House School was so arranged as to contain a great deal of material and the adults helped the children consistently to find solutions to the problems which they set for themselves. The children asked questions and initiated activities; but so also did the adults. The art consisted in initiating an interest or an activity which seemed to be already latent in the group. The role of the teacher, therefore, is certainly to guide, through the provision of materials, through directing attention to problems and processes likely to interest children, through asking appropriate questions and through tactful help at critical points in finding solutions or in carrying out tasks which may be within the conception of the pupils but beyond their immediate ability to execute.

Praise and Blame

The maintenance of interest and curiosity, however, at least in part, depends upon success and acceptance and praise that effort brings. Some light is thrown on this by an early experiment by Hurlock. She took four groups of children, all of whom were set the task of addition. Initially, all groups achieved approximately twelve points of score on the test used. One group was then praised for its performance; a second was reproved

for the mistakes which it had made; a third was ignored, although the pupils heard the praise and reproof meted out to the others; finally, there was a control group. After a period of this treatment, the four groups were retested with the results that: the praised group had a mean of twenty points; the reproved group fourteen points; the ignored group twelve points; and the control group eleven points.

This experiment suggests that praise is a powerful stimulus to effort at least with primary school children. It also seems to be clear that blame or punishment of whatever kind has the effect of inhibiting a particular form of action without necessarily putting anything positive in its place. The effects of blame seem to be primarily emotional, affecting performance rather than learning and in situations where what is wanted is unclear or beyond the capacity of the pupil it may lead to the repetition of undesirable behaviour or to confusion. The relationship of this last observation to some of the more stubborn cases of arithmetic or reading failure is evident.

Lewin also points out that a reward or punishment situation, particularly where it is directed to individuals in the class group, must imply some form of policing; that is to say, that barriers are erected against the attainment of the reward by a short cut such as copying the answers. Such externally imposed frustrations are bound, he suggests, to have a profound effect upon the social group concerned.

The effectiveness of praise or blame seems to be conditioned by other factors. A research by Thompson and Hannicutt with pupils in the Vth grade (i.e. about 11-12 years old) suggested that repeated praise tended to increase the work of those children who were more introverted while the more extroverted children were more stimulated by blame.

Still more is it likely to be conditioned by the prior experience of pupils, especially in so far as they have been successful or unsuccessful in similar activities before. Three researches are of interest here, carried out by Sears, by Gebhard and by Cartwright. Sears found that pupils with a prior experience of success tend to be realistic in their goal-setting in arithmetic and reading, that is to say that their level of aspiration more or less coincides with their possibilities. On the other hand those who have experienced failure in similar tasks, either set unrealistically high goals or they tend to state a much lower level of aspiration as a kind of self-protection against further failure. Cartwright and Gebhard, working independently, discovered that success and the anticipation of success increased considerably the attractiveness of otherwise neutral tasks.

The effects on children of praise and blame, success and failure are thus difficult to predict in any given case. It is evident that failure in a general context of success has a different meaning from a failure which is one more of a long chain. It is evident, too, that the degree to which the child feels himself to be generally acceptable to his teacher will condition markedly the effect of blame or failure on his attitudes to further effort.

Anxiety

There is some evidence that a certain amount of anxiety in otherwise normal children provides a motivation for learning. Robertson, in an unpublished research on children whose performances were better or worse in the secondary grammar school than was

predicted by the eleven plus allocation tests, finds some support for the idea that children who improve are those who are mildly anxious about their work. Too great an anxiety on the other hand (whether about school matters or about matters in the out-of-school life), may lead to inhibition of effort or passivity. Equally, it seems that children who are without anxiety do not do as well as might be expected. This last finding coincides with other work that suggests that the ablest children in a group tend to conform to the norm rather than too markedly to excel. Lynn, studying the reading and arithmetic attainments of primary and secondary school children, found that anxiety tended to be associated with a higher level in reading than in arithmetic in otherwise normal groups.

A level of anxiety which is noticeable to the teacher may, however, tend to inhibit and disorganise effort—particularly effort in such subjects as arithmetic which depend upon undisturbed concentration. Many normal educational circumstances, notably examinations, provoke evident anxiety in children and adolescents and this may in fact adversely affect performance even of normal pupils. Calvin, McGuignan and Sullivan, working with undergraduate women students, found that examination performance was markedly improved if an outlet for such anxiety (in the form of free comment) was permitted.

Work of this kind, and general psychological considerations, suggest that the teacher should stimulate children to effort and achievement—perhaps rather more than is customary in some progressive schools and activity classes—but that such stimulation should be adjusted by an understanding of the individual pupil's general proneness to anxiety and insecurity. Furthermore, it seems likely that freedom on the part of the pupils to express their anxieties coupled with skilful reassurance on the part of the teacher is likely to ensure that such anxiety is in fact creative rather than inhibiting.

Group Influences

So far, we have spoken of the child as an individual or as one side of the teacher-pupil relationship; he is, however, much more than this. At least when he comes to school and often before, he becomes aware of the group of his contemporaries and equals and slowly begins to develop a group loyalty. Indeed, the feeling of belonging either to a small group of friends or to a larger class-group is an important element in his feeling of security; and remaining acceptable to the group becomes itself an important motivation for him. Within the group, too, he has a role to sustain which is partly a matter of his own choice and partly what the others reflect upon him. Thus, by middle childhood at least, what might be called socially determined motivations are of great importance inside and outside the classroom and the simple, direct relationship of teacher and taught no longer quite holds. The aims of the teacher and those of the taught may be in conflict without either side being fully aware of why, or even that this is so.

Some of the work which has been done on the relationships between the personality and methods of the teacher and the reactions within the class is highly relevant here. Anderson and his collaborators, working with nursery school classes, found that the teacher who tended to dominate provoked resistance, fidgeting and over-conformity in her pupils. In contrast the teacher who attempted integrative responses to children and who sought to draw them into activities with their own free consent was rewarded by

more spontaneity and more initiative in learning from her class. Lewin, Lippit and White, working with somewhat older children and with adolescents in youth clubs, contrasted three broad kinds of approach: dictatorial; laissez-faire; and democratic leadership. They found that hostile and aggressive reactions were as much as thirty times more frequent in autocratic groups. In the groups which were democratically led there was higher morale, more communication, closer friendships, less discontent and aggressive behaviour, and more work-minded conversation. Similar conclusions are reached by Richardson in her studies of teacher behaviour and adolescent reactions.

Co-operation and Competition

Closely related to this are experiments which have been conducted to determine whether co-operative working or interpersonal competition gives the better results. Deutsch, in a study of two matched groups of school children, found that competition resulted in much more destructive and aggressive criticism while the co-operative group showed more co-ordination of effort, greater pressure to achieve, more productivity per unit of time, and a higher quality of the end-product. In addition there were advantages of a social kind shown in friendliness and attentiveness to others.

Thelen, summarising the upshot of a long series of researches carried out in the classroom laboratory of the University of Chicago, puts forward ' some significant propositions ' as follows: (*a*) the teacher's behaviour in a large measure determines the quality of the emotional conditions in the classroom; (*b*) the learning of certain social attitudes and human relations principles is affected by teacher-pupil interaction; (*c*) the teacher-pupil interaction patterns may affect the pupil at deep (i.e. subconscious) levels; (*d*) pupil-pupil interaction (for example, status-roles) can be affected by the nature of the teacher-pupil interaction. He further points out that much of this can be generalised from school to adult groups; conversely, of course, much of the work on decision-making in industry is applicable to school situations and it has been over and over again found that a genuinely democratic participating group tends to produce in each member more satisfaction in the task undertaken and a tendency to retain the group decision as one's own.

A word of caution, however, is necessary in attempting to apply these findings. It does seem clear that in general the best motivation is obtained when the teacher involves the class in the making of the decisions as to what shall be done and how. Perhaps the best and almost the sole large-scale demonstration of the effectiveness of this is the well-known Eight Year Study in which the best results were obtained from those schools which approached their work in such a way as to obtain the maximum participation of the students.

However, although accepting and democratic, the teacher should not abandon his necessary role of leadership. He has a contribution to make to the decision-making and to the choice and setting of aims which springs from his superior knowledge and experience. He must, too, allow for the fact that not all pupils will have the same willingness or ability to participate and that some at least may well be isolates who do not respond at all or at least directly or desirably to group influences. He has to guard against the tendency of the abler members to adapt their effort and pace to that of the group rather than to exert themselves to the full—thus, he may have in fact to combat an attitude, itself the

product of group loyalty, that the individual should not differ too much from his companions. Equally, he has to be aware of the tendency of some children to become passengers.

He should, too, know that he cannot change the climate of a group abruptly from the more usual one of more or less benevolent autocracy to a truly democratic one; the change has to be mediated and the necessary attitudes have in fact to be learned by all who participate, including, very often, the teacher himself. Finally, the nature of the leadership which can be given by the teacher will depend considerably upon the degree of maturity of the pupils and upon the realism with which they can set goals for themselves. In general, the older and more able the group of pupils, the more fully democratic the leadership can be.

Socio-economic Status

We have so far spoken as though the motivations for learning were either individual or a function of the nature and climate of the classroom group. This, of course, is only partly true. Outside these and beyond the control of the teacher lie certain social motivations, difficult to analyse, but which have to be taken into account. Our knowledge of these is sketchy and indirect, and the interpretation of the data which we have is somewhat speculative. There is some evidence, both from English and from American sources, to suggest that general interest and success in school work are positively related to socio-economic status, even when general ability is held constant. The interest of the parents in their children—an influence beginning well before school years—their own experience with schooling and their degree of education seem also to be directly connected with levels of attainment. On the other hand, there may be, in our community, a rural group and a small urban one also which is indifferent if not hostile to education.

Attitudes and Opinions

The influence of the social group operates in other ways. It is well known that the attitudes and opinions of children and adolescents tend to be closer to those of their parents than to those of their teachers and that the vocational choices which are made by boys and by girls tend in many cases to be related to the occupation of the parents. An enquiry by Jahoda carries this into more detail. He studied school-leavers from secondary modern schools and found that whatever their actual social class membership, boys tended to identify themselves with the skilled manual and routine grades of non-manual work. They intended to ' stay with the lads ' and declared that they wished to become fully skilled craftsmen, which was what a ' job with prospects ' implied.

The girls tended to be less realistic in their assessment of the present status of their parents and in their ambitions; many of them, too, seemed to express a conflict between their desire to move up in social grade and loyalty to the group to which they are attached. In the ' teens ', and especially in the last year or two of compulsory schooling, such vocational identifications may be important sources of motivation; but they do not necessarily operate in the direction proposed by the teacher.

Conclusion

It is apparent from the analysis we have made of some of the work on children's motivations that no simple rules and precepts can be given to teachers. It is, too, very clear that psychologists are far from having unravelled the whole delicate system which determines whether or not a class will put forth a maximum effort in school. However, such knowledge as we do have can lead us to a number of practical inferences which, if they guide classroom practice, are likely to be helpful.

The first of these is embodied in the proposition that if the fundamental needs of the growing personality are not fully met, there will be little energy left for intellectual things. Some, even most, of the dissatisfaction and maladjustments of children and adolescents arise outside the school. This, however, does not mean that the school is powerless to do anything about them. Severely maladjusted children should, of course, be referred for examination and treatment to the school psychological service; but all children at some stage show signs of strains in their growth which may affect their effort in class. It is then that they specially demand perceptive handling and support from their teachers and parents.

With advancing maturity and emotional balance, children normally learn to differentiate the demand for love which they make on their parents from that they make of their teachers and friends. But in spite of his deceptively self-contained air the primary school child still needs affection and acceptance in school; and so less obviously perhaps but none the less imperiously does the adolescent. An impersonal or worse still, a hostile attitude on the part of the teacher, or the school staff generally, is not conducive to effort at any stage. Warmth and understanding from the teacher are essential to good pupil-teacher relationships and it is in the context of these that praise and blame, punishment, reward and criticism must be viewed as stimuli to effort.

Children of all ages have other needs which should not be denied. Some outlet for aggression, some chance of physical movement, the security which comes from an orderly and comprehensible framework of do's and don'ts and from a reasonable but not obsessive rhythm to the school day, are likely to allow the secondary motivations of curiosity and the desire to explore and manipulate to emerge. So, too, self-assertion, both as an individual and as member of a group or team, calls for the acceptable and constructive outlet of success and a place in the sun.

Play and fun, too, are fundamental to growth and not merely to the young child. Through play and fantasy children and adolescents can come to grips with and solve some of their problems. Many children in our society do not have sufficient opportunity to play in their out-of-school lives and suffer in consequence. A major psychological argument for carefully planned active methods is that among other things, they allow for those who are deprived of play.

Some work by Ojemann and by Selley has drawn attention to the influence on children's school progress of greater knowledge by their teachers of childhood and its problems and of the value to normal school work of the chance given to primary school pupils of discussing their personal problems freely. This does not imply an airy-fairy 'psychiatrisation' of the school. It suggests that when outlets are given for pressing

needs and when the teachers thoroughly understand the emotional growth of children, then their disciplinary systems, and the incentives which they use to increase motivation are more likely to be effective and less likely to run counter to the deep and powerful drives and preoccupations of their pupils.

Fear and anxiety have been shown to be, on the whole, inhibitive and disorganizing factors in learning. Some children it is true do seem to thrive under punishment, but in the long run, it seems more likely to lead to undesirable forms of social learning—cheating, deception, and so on—than to any good effects. This, however, does not mean that punishment, reproof and blame should never be used. In the context of healthy pupil-teacher relationships and when they are clearly deserved, they are likely to be effective as deterrents. However, this effect is largely lost if at the same time pupils are not constructively shown the way to win praise and approval.

It is evident, too, that children's levels of aspiration, the goals they set themselves, are related to two things which are within the control of the school. Prior knowledge and learning and especially the experience of successful effort, affect incidental learning and present motivation. Thus, the adjustment of the gradient of difficulty in all tasks which pupils undertake is vital not merely for the matter in hand but for the growth of attitudes towards work and learning in the future. This does not mean that work should be so arranged that success comes easily. It means that the demands made are such as to evoke the maximum effort and that the teacher should be watchful to see that failure is only a rare occurrence; he should intervene to help when the child is clearly struggling with something beyond his powers or will to succeed. Here, too, we should point out the value of positive habits of getting down to a task in the expectation that an honest effort will bring success or at least understanding from the teacher.

Finally, the importance of the class as a group and of the sub-groups within the class must not be overlooked and it is in drawing attention to this that perhaps the greatest contribution has been made by recent research. The system of discipline used by the teacher, his personality and the relationships he builds with individual pupils all affect the climate and consequently the motivation of the class. He is most likely to succeed if he becomes aware of the feelings of group solidarity which exist, if he is perceptive of those children who tend to be isolated and if he attempts democratically to ensure that the pupils participate in the determination of the immediate, and with older children and adolescents, the more remote aims of their own education. This does not mean that he abrogates his duty to give direction, to set standards, to maintain discipline; it means that he exercises it in a way which is intelligible and acceptable to his pupils, and in a framework in which each child is sure of being cared for. Even more important perhaps is that in such a social climate he is more likely to be able realistically to assess what his pupils can achieve and help them, through the experience of success, to develop the habit and the will to learn.

Bibliography

This is a list of the material consulted and summarized in the preceding paper. Items providing a select initial bibliography for further reading are marked with one asterisk (*); more difficult and specialised items with two asterisks (**).

ANDERSON, H. H., *et al.*, " Studies of Teachers' Classroom Personalities." *Appl. Psychol. Monogr.*, 1945-6, Nos. 6, 8, 11.

AUBLE, D., and MFCH, E., " Response Strength in a Classroom Task Related to a ' Forward ' Delay in Reinforcement." *J. Ed. Psych.*, Vol. XLV, No. 3, March, 1954.
BROWN, W. F., ABELES, N., and ISCOE, I., " Motivational Differences between High and Low Scholarship College Students." *J. Ed. Psych.*, Vol. XLV, No. 4, April, 1954.
*BUGELSKI, B. R., *The Psychology of Learning*, Chap. 9, " Motivation and Learning." London: Methuen, 1956.
CALVIN, A. D., MCGUIGAN, F. J., and SULLIVAN, M. W., " A Further Investigation of the Relationship between Anxiety and Classroom Examination Performance." *J. Ed. Psych.*, Vol. XLVIII, No. 4, April, 1957.
CAMPBELL, W. J., " The Influence of Home Environment on the Educational Progress of Selective Secondary School Children." *Brit. J. Ed. Psych.*, Vol. XXII, Part II, June, 1952.
*CARROLL, HERBERT A., " Motivation and Learning: Their Significance in a Mental-Health Program for Education." *Mental Health in Education*, 54th Yearbook of the U.S.A. National Society for the Study of Education, Part II. Univ. of Chicago Press, 1955.
CARTWRIGHT, D., " The Effect of Interruption, Completion and Failure upon the Attractiveness of Activities." *J. Exp. Psych.*, Vol. XXXII, 1942.
CASSELL, RUSSELL N., " Motivation as a Synthesis of Contemporary Psychology." *J. Ed. Psych.*, Vol. XLIII, No. 3, March, 1952.
COLLINS, M., " The Causes of Premature Leaving from Grammar Schools, Part II." *Brit. J. Ed. Psych.*, Vol. XXV, Part I, February, 1955.
CRAIG, R. C., " Directed versus Independent Discovery of Established Relations." *J. Ed. Psych.*, Vol. XLVII, No. 4, April, 1956.
*DANIEL, M. V., *Activity in the Primary School*, Oxford, Blackwell, 1947.
DEUTSCH, M., " An Experimental Study of the Effects of Co-operation and Competition upon Group Process." *Human Relations*, Vol. II, No. 2, 1949.
DREYER, A. S., " Aspiration Behaviour as Influenced by Expectation and Group Comparison." *Human Relations*, Vol. VII, No. 2, 1954.
**EYSENCK, H. J., *The Structure of Human Personality*, Chap. VI, " The Analysis of Physiological Measures " London: Methuen, 1953.
*FLEMING, C. M., " The Child Within the Group—The Bearing of Field Theory and Sociometry on Children's Classroom Behaviour." *Studies in Education, No. 7: The Bearings of Recent Advances in Psychology on Educational Problems*. London: Evans Bros., 1955.
*GARDNER, D. E. M., *Long Term Results of Infant School Methods*. London: Methuen, 1950.
GARDNER, D. E. M., *Testing Results in the Infant School*. London: Methuen, 1942.
GEBHARD, M. E., " The Effect of Success and Failure upon the Attractiveness of Activities as a Function of Experience." *J. Exp. Psych.*, Vol. XXXVIII, 1948.
HARLOW, HARRY F., " Motivational Forces Underlying Learning." In *Learning Theory, Personality Theory and Clinical Research*. The Kentucky Symposium. Lonson: Chapman & Hall, 1954.
**HEBB, D. O., *The Organization of Behavior*, Chap. 8, " Problems of Motivation—Pain and Hunger." London: Chapman & Hall, 1949.
HIERONYMUS, A. N., " A Study of Social Class Motivation: Relationships between Anxiety for Education and Certain Socio-Economic and Intellectual Variables." *J. Ed. Psych.*, Vol. XLII, No. 4, April, 1951.
*HIGHFIELD, M. E., and PINSENT, A., *A Survey of Rewards and Punishments in Schools*. Publication No. 3 of the National Foundation for Educational Research. London: Newnes, 1952.
HILGARD, ERNEST R., " Human Motives and the Concept of the Self." *The Study of Personality*. London: Chapman & Hall, 1954.
*HILGARD, E. R., and RUSSELL, D. H., " Motivation in School Learning "; SWENSON, E. J., " Applications of Learning Principles to the Improvement of Teaching in the Early Elementary Grades "; THELEN, H. A., and TYLER, R. W., " Implications for Improving Instruction in the High School." *Learning and Instruction*, 49th Yearbook of the U.S.A. National Society for the Study of Education, Part I. Univ. of Chicago Press, 1950.
HUGHES, E. W., " Children's Choices in Individual Activities in the Junior School." *Brit. J. Ed. Psych.*, Vol. XXV, Part I, February, 1955.
HURLOCK, E. B., " An Evaluation of Certain Incentives Used in School Work." *J. Ed. Psych.*, Vol. XVI, 1925.
**ISAACS, S., *Intellectual Growth in Young Children*. London: George Routledge & Sons, 1930.
JAHODA, G., " Social Class Attitudes and Levels of Occupational Aspiration in Secondary Modern School Leavers." *Brit. J. Psych.*, Vol. XLIV, Part 2, May, 1953.
KEMP, LESLIE C. D., " Environmental and Other Characteristics Determining Attainment in Primary Schools." *Brit. J. Ed. Psych.*, Vol. XXV, Part II, June, 1955.
KITTELL, JACK E., " An Experimental Study of the Effect of External Direction during Learning on Transfer and Retention of Principles." *J. Ed. Psych.*, Vol. XLVIII, No. 7, November, 1957.
*KNIGHT, REX, " Children's Needs and Interests—Contemporary Theories of Motivation." *Studies in Education, No. 7: The Bearings of Recent Advances in Psychology on Educational Problems*. London: Evans Bros., 1955.
*LAIRD, A. J., and KNIGHT, A. R., " Contemporary Studies of Motivation." In *Current Trends in British Psychology* London: Methuen, 1953.
LEWIN, K., " Field Theory and Learning." In *The Psychology of Learning*, 41st Yearbook of the U.S.A. National Society for the Study of Education, Part II. Univ. of Chicago Press, 1942.
LIPPIT, R., and WHITE, R. K., " The Social Climate of Children's Groups." Chapter XVIII in Barker, Komim and Wright (Ed.), *Child Behaviour and Development*. New York: McGraw-Hill, 1943.

LYNN, R., " Temperamental Characteristics Related to Disparity of Attainment in Reading and Arithmetic. ' *Brit. J. Ed. Psych.*, Vol. XXVII, Part I, February, 1957.
MCGEOGH, J. A., and IRION, A. L., " Learning as a Function of Motive-Incentive Conditions." In *The Psychology of Human Learning*. New York: Longmans, Green & Co., 1952.
MORGAN, M. I., and OJEMANN, R. H., " The Effect of a Learning Programme Designed to Assist Youth in an Understanding of Behaviour and its Development." *Child Development*, Vol. 13, No. 3, 1942.
*MORRIS, BEN, " Mental Health in the Classroom." *Studies in Education, No. 7: The Bearings of Recent Advances in Psychology on Educational Problems*. London: Evans Bros., 1955.
* JESER, O. A. (Ed.), *Teacher, Pupil and Task—Elements of Social Psychology applied to Education*. London: Tavistock Publications, 1955.
OJEMANN, R. H., and WILKINSON, F. R., " The Effect on Pupil Growth of an Increase in Teachers' Understanding of Pupil Behaviour." *J. Exp. Educ.*, Vol. 8, December, 1939.
RADKE-YARROW, M., and YARROW, L. J., " Child Psychology." *Annual Review of Psychology*. Stanford, California-Annual Reviews Inc., Vol. 6, 1955.
**RICHARDSON, J. E., *et al.*, *Studies in the Social Psychology of Adolescence*. London: Routledge and Kegan Paul, 1951.
SAWIN, E. I., " Application of Learning Theory to Appraisal of Motivation: A Pilot Study." *J. Educ. Res.*, Vol. XLIX, No. 8, April, 1956.
SEARS, P. S., " Levels of Aspiration in Academically Successful and Unsuccessful School Children. *J. Abn. and Soc. Psych.*, Vol. XXXV, 1940.
SEELEY, J. R., *Crestwood Heights*. Toronto: 1956.
**SHERIF, M., and CANTRIL, H., *The Psychology of Ego-Involvement*. London: Chapman & Hall, 1947.
SILBERMAN, H. F., " Effects of Praise and Reproof on Reading Growth in a Non-Laboratory Classroom Setting." *J. Educ. Psych.*, Vol. XLVIII, No. 4, April, 1957.
SIMPSON, R. H., " An Evaluation of Motivation, Assignment Making, Problem Procedures, and ' Service Learnings ' in a School System." *J. Educ. Res.*, Vol. XLIV, No. 1, September, 1950.
STILES, F. S., " Developing an Understanding of Behaviour at the Elementary School Level." *J. Educ. Res.*, Vol. XLIII, 1950.
THELEN, WITHALL, FLANDERS, REHAGE, PERKINS, GLIDEWELL, Symposium on " Experimental Research Towards a Theory of Instruction." *J. Educ. Res.*, Vol. XLV, No. 2, October, 1951.
THOMPSON, C. G., and HUNNICUTT, C. W., " Effect of Repeated Praise or Blame on the work achievements of Introverts and Extroverts." *J. Educ. Psych.*, Vol. XXXV, May, 1944.

Teaching Children to Read—1

THE RELATIVE EFFECTIVENESS OF DIFFERENT METHODS OF TEACHING

A. The Place and Value of Phonics

by JOYCE M. MORRIS

THE effectiveness of any teaching method is determined by school conditions as well as by the attributes of teacher and pupil. It is therefore unlikely that research will ever reveal one best method of teaching reading but, when sufficiently developed, it may help the inexperienced teacher to choose suitable reading methods in a given set of circumstances, and thereby avoid much wasteful experimentation in the classroom.

Basic Methods of Teaching Reading

Today selection may be made from a great number of reading systems of which almost all are modifications and combinations of a few basic methods. Two of these—the alphabet-spelling method which originated in ancient times, and the look-and-say or word method devised by Comenius in the 17th century, were almost universally employed in our early infant schools. By these methods children learning to read the word CAT, for example, either named the separate letters as SEE AY TEE and followed this by saying CAT, or were taught to recognise the word correctly at sight without more ado.

The alphabetic method is undoubtedly of value in learning to spell but of limited usefulness for the reading of words and sentences; the look-and-say method, although based on the psychologically correct idea that a word is not merely a collection of letters but has a recognisable character or pattern of its own, does not by itself give pupils a means of tackling new words independently. Hence, it is not surprising that these early methods lost ground rapidly when phonic methods were widely introduced and developed. With the substitution of sounds for letter-names and their subsequent blending to form a whole word, the word CAT, for instance, could now be attacked as follows: KUH-A-TT, CAT. This type of blending method is still used, especially by the older generation of teachers, but in modern phonic teaching it is more usual to find children being taught to divide simple words into initial consonants and phonograms. Thus, the word LAND would be divided into L-AND, and then by combining the phonogram with different preceding and following letters a word family may be built up, e.g., band, stand, sandy, etc. Training in the recognition of common syllables and in analysing long words by syllables is now a recognised part of phonic training.

Phonic methods help children to learn to read in a systematic left-to-right direction and provide them with a technique by which unfamiliar words may be deciphered. However, apart from the fact that a considerable proportion of English words are non-phonetic, other disadvantages of these methods are frequently quoted. It is claimed that the mechanical nature of any phonic method lays too little stress on the importance

of comprehension in reading, and that the practice of splitting up words into their common sounds or combinations of sound hinders whole word recognition. The phonic method imposes restrictions on the vocabularies of reading primers so that they are often poor in content and artificial in style. Habits of slow, laboured reading are also said to be developed by this method.

Different phonic systems flourished at the beginning of this century, but, when the limitations of a purely phonic approach were realised, teachers turned again to the older methods. The highly mechanical, difficult and uninteresting alphabetic method was never reinstated to its former position in the reading programme, but the value of a sight vocabulary was again recognised with the result that a ' mixed ' method was recommended whereby regular words were taught phonically and irregular words by look-and-say.

Fresh support was also given to advocates of ' analytic ' as opposed to ' synthetic ' methods when the sentence method of teaching reading was popularised in the thirties. In this method (which is claimed to be a good one because it follows natural lines), the sentence, rather than the word, is regarded as the unit of meaning and emphasis is laid on reading comprehension. Its great value lies in the help offered by the context which, although limited to a certain extent by the need for frequent vocabulary repetition, does not have to be of the cat-sat-on-a-mat variety. Children taught exclusively in this way, however, may be satisfied with guesses instead of looking carefully at individual words, and they will probably be bewildered when they meet a word which is new to them. Difficulty in distinguishing between words of similar size and shape may also be experienced.

Since satisfactory progress is not guaranteed by the use of these visual and auditory techniques alone, many teachers include kinaesthetic methods in their reading programmes. These methods, which stress the importance of tracing and writing in learning to read, were first described over two thousand years ago. More recently, Montessori employed them, but they were not recognised widely as basic procedures until 1921 when Fernald and Keller reported their favourable effects on the development of word recognition in non-readers.

Kinaesthetic methods reinforce other sensory impressions of words, develop a consistent left-to-right direction in reading, and provide adequate drill and repetition. They also encourage careful and systematic observation and enable both teacher and pupil to measure progress easily. Three major criticisms of the exclusive use of these methods may be offered. First, they do not give pupils a means of deciphering new words without assistance. Secondly, they develop certain movements of arm, hand and lips which may interfere with later progress in speed and fluency. Finally, the necessarily individualised instruction which these methods entail make them generally more suitable for small classes.

There are thus advantages and disadvantages in all the main methods of teaching reading used in the past. Consequently, most teachers today do not confine themselves to one particular method, but, recognising the importance of individual differences, provide sufficient variety in their approach to reading to enable their pupils to read by the look of a word, by its sound, by guessing from the context, and by utilising the sense of

movement. They differ, however, in the proportions of each method which make up their basic reading programme, and in the extent to which they adjust this proportion to the needs of individual children. There are also differences in the emphasis which teachers place on oral and silent reading as well as in the amount of systematic as opposed to informal instruction given. It is at this point that controversies arise, and where scientific investigation may be of practical use in deciding what is empirically sound educational practice and what is merely a matter of ' fashion '.

Reading Methods and Research

Since the second decade of this century experiments to determine the relative effectiveness of different teaching methods have been carried out. Unfortunately, few investigators have exercised sufficient care in ensuring that differences in results have been due to teaching method and not to the skill of the teachers, the attributes of the children or other school characteristics. Methods, too, have not been clearly defined, and uniformity of name does not necessarily guarantee uniformity of method. Thus, seemingly conflicting evidence has been produced on many controversial questions.

Several attempts to draw together what is certainly known about reading in general have been made, among the most notable being those of Gray, Vernon, and Traxler and Townsend. Recent events, however, have focussed attention specifically on reading methods, and on the value of phonic as opposed to whole-word methods in particular. In 1954, public concern about school methods was aroused in America by the publication of a book by Flesch addressed to parents and entitled *Why Johnny Can't Read.* The author claimed that there had been a deterioration in children's reading standards which was largely due to the neglect of phonic teaching in the modern school. Flesch also maintained that current educational practice is not in line with scientific evidence, since he can find none in favour of the whole-word approach although he is able to quote a number of studies which prove the superiority of the phonic method. A similar conclusion was reached by Daniels and Diack in this country, for they stated that: "In all the experiments reported in the literature, we know of only three which have not come to the definite conclusion that some variety of the phonic method gives better results than a whole-word method as far as word recognition is concerned." According to these researchers even this small amount of evidence is suspect owing to the experimental design of the three investigations cited.

These statements and the concern of parents and teachers naturally call for a re-examination of modern methods of teaching reading and the scientific bases on which they rest. Since, however, the literature on reading methods is extensive, the present discussion will be confined to one main topic only, namely, the *place and value of phonics in the reading programme.*

Consideration of what research tells us about phonics necessarily includes references to other methods, but the major purpose of this first article is to review studies which help to answer such questions as: Should phonics be taught? When and how should phonic training be given? Is a particular kind of reading skill produced by systematic phonic instruction? Are phonic methods more effective with certain types of children, e.g., the

dull and backward? First it is necessary to meet the challenge that phonics teaching is neglected in the schools of today.

The Place of Phonic Methods in Current Educational Practice

Before the book by Flesch was published, 220 experienced teachers and other school people from 33 different States of America were asked by Russell what place they actually gave to phonics in their teaching of reading. According to the replies, 15 per cent. gave phonics instruction in lessons regularly scheduled but separate from other reading lessons; 4 per cent. taught phonics in all reading lessons; 39 per cent. in most reading lessons; and a further 39 per cent. as part of a few reading lessons. Only 2 per cent. of the teachers omitted phonics training altogether.

Further questions revealed that whilst in the sample studied, there is no doubt about phonics being used, it is equally clear that they are regarded as only one means of developing word-attack skills, which in turn are viewed as only one goal in the total reading programme.

Similarly, when the present writer recently visited a random sample of 60 Kent primary schools, every school was found to be giving some phonic training in conjunction with other methods, but with different emphasis according to the children's needs and the beliefs of their teachers. In 13 schools children began learning to read by phonic methods, and in six of this group they were only five years old when introduced to them. Moreover, the head teachers of two-thirds of the 60 schools stressed that backward readers in the transition classes and where necessary further up the school, were receiving a great deal of phonic instruction.

There is thus some evidence that phonic skills are being taught in schools at the present time, although prolonged classroom observation in a large and representative sample of schools is necessary before the true position of phonics may be ascertained.

Granted that most teachers use phonic methods today it does not necessarily follow that they should do so. It will be remembered that the value of the alphabetic method for spelling blinded teachers for centuries to its relative ineffectiveness as a method of teaching reading.

The Value of Phonic Methods

Phonic methods have long been the subject of controversy and many investigations have been carried out in an attempt to establish their value.

One of the earliest studies was that of Gill who, in 1912, reported that he had compared the reading ability of three groups of top infants from schools using three different methods, namely, the ' Dale ' method of phonic instruction, a phonetic method in which the forms of the letters had been made attractive by ingenious but irrelevant associations, and the sentence method. The first two methods had been used for two years and the third for sixteen months. Reading ability was measured by the time required to read two passages taken from an infant primer but printed without punctuation or spacing between the words, and in one case with the initial letters of nouns printed in capitals. Comparing the average times, the fluency and intelligence of reading of the

three groups, Gill concluded that the sentence or whole-word method was of greater practical value than either of the two phonic methods.

Since Gill's report was published at a time when phonic methods were still in the ascendancy his findings were rather disconcerting. However, a year later, in a paper reporting his own experiments, Valentine stated that little significance could be attached to the results of this study since the comparative intelligence of the children was unknown, the general efficiency of the teachers had not been considered, and the form in which the passages were printed favoured the group which had learned to read by sentence wholes. Later investigators have also pointed out that the sentence method group was further favoured by the use of speed as the criterion of reading ability, and by the fact that children in this group were older and presumably more ' ready ' for instruction when it was given.

In his own research, Valentine compared two groups of children, aged six-and-a-half to eight years, in their ability to read English written by means of a Greek alphabet after having been taught by a phonic method and a look-and-say approach. He concluded that:

> " There seems to be nothing so inherently difficult, even for little children, in the synthetic work involved by the Phonic method as has been asserted by some of its critics. . . . Children taught by the Phonic method do better than those taught by the ' Look and Say ' method both in reading words previously seen and words previously unseen, and even in some words seen by the ' Look and Say ' children but not seen by the Phonic children. There seems to be some evidence, however, that for very dull children the ' Look and Say ' method is more efficient."

An experiment involving the teaching of Greek letters to children with some experience of the English alphabet may be criticised as artificial but at least Valentine made sure of the initial comparability of his samples. Furthermore, by drawing attention to the importance of experimental design and to the fact that a given method might not be equally effective with all pupils, this investigation made a valuable contribution to further studies as well as providing evidence favourable to phonic methods.

It would be impossible, even if it were desirable, to describe the numerous inquiries provoked by the conflicting results of Gill and Valentine. Mention should be made, however, of a frequently quoted experiment by Winch who, in a period when whole-word methods were gaining ground in the schools, produced further evidence of the value of phonic methods. For this experiment, designed to determine the relative effectiveness of the alphabetic, phonic and look-and-say methods as well as a special phonic method called ' phonoscript ', two groups of 32 children between five and six from a school in a poor district of London were chosen. They were equated on the basis of preliminary tests of ability to learn associations between shapes and sounds, and then given twenty-five formal reading lessons. At the end of this period the children were tested individually for their ability to read material consisting of words they had already learnt and a certain number of new words.

Results favourable to phonics and to the more recent system of ' phonoscript ' in particular were obtained. Perhaps of even greater interest to teachers, however, was

Winch's final recommendation that, " English children under our conditions of school organisation, may quite profitably begin definitely to learn to read at an average age of 5 years 3 months. . . . In schools where children are well born and of good environment they may begin earlier; in schools attended by very poor children, later." Conclusive evidence of the relationship between socio-economic background and reading ability has since been obtained, but this is one of the first occasions when a research worker pointed to the necessity of adapting teaching methods to the home environment.

The way in which this inquiry was conducted has been praised highly, but it might be said that twenty-five lessons spread over a two-month period were too few to give a reliable comparison of methods, and that, if the lessons had continued, the initial advantage of the phonic methods might have been lost to the more interesting whole-word approach. Further, as in all experiments where one teacher is responsible for all of the teaching, individual differences in the skilful use of a particular method must be taken into account when assessing the results.

In the period following Winch's investigation most of the experimental work on reading methods was carried out in the U.S.A. There, Mosher and Newhall reported in 1930 that there was no essential difference between the effects of a phonic and word approach at least at the end of the second school year. One might question this finding, however, in that accuracy in word recognition was not tested in their investigation and this is the aspect of reading skill to which phonics is supposed to contribute most directly. On the whole, the verdict of American research seems to be in accord with Valentine and Winch in finding that phonic methods are of value, although one authority, Gates, has stated that if it were necessary to choose between phonic and non-phonic training the latter would be selected.

In recent years, the problem has become not the simple black and white one of phonics versus no phonics but rather the more complex one of when, how much, for what purpose, and to whom phonics should be taught. A final verdict on some of these questions has not yet been obtained, but consideration of a few of the more frequently quoted studies may give some idea of the present state of our knowledge about phonic methods. Only reference to researches carried out in English-speaking countries will be made since the effectiveness of phonic methods is obviously determined to a large extent by the nature of the language concerned.

The need for phonic readiness

When experiments on reading methods were first conducted, it was customary to start all children with phonics from the beginning; in current practice phonic methods are not usually introduced formally until a later stage in the infant school course. How far research favours this practice may be seen from the following investigations.*

In 1925, nearly one thousand pupils in eight schools of various types from different sections of Newark, New Jersey, were divided into phonic and non-phonic instructional groups. Their teachers were required to alternate between the methods in an attempt

*In evaluating the evidence of research carried out in other countries it is important to be aware of differences in school organisation. For example, only about 50 per cent. of five-year-old children in the U.S.A. are enrolled in school compared with the total child population in this country. American elementary schools usually have either six of eight grades and, where kindergartens are not maintained, children enter the first grade at about the age of six years.

to obviate the effect of differences in teaching efficiency. At the end of five months of training and again at the end of ten months, the children were given a number of reading tests and the results for those pupils who had received uninterrupted instruction were compared. Further tests were also made on 426 pupils who continued in controlled groups through the first half of the second grade.

According to the experimenters, Sexton and Herron, the results clearly indicated that the teaching of phonics functions very little or not at all with beginners in reading during the first five months, but begins to be of some value during the second five months, and is of greater value in the second grade. There was also evidence that the teacher is more important than the method, for where results were unusually good in a class taught by a teacher using phonics they were also unusually good when the same teacher taught without phonics. On the other hand, poor results were obtained in both phonic and non-phonic groups taught by the same teacher.

A few years later, Garrison and Heard carried out a three-year sequential study of 111 first-graders, divided into four groups on the basis of intelligence test scores. In four different classrooms with approximately equal school surroundings and quality of teaching, two groups received phonic instruction and two had none.

These investigators came to the conclusion that phonic training should be deferred until the second or third grade (i.e., pupils aged from 7-8 years, approximately). They also reported that whilst instruction in phonics develops independence in word pronunciation, children with no phonic training make smoother and better oral readers in the lower grades. Moreover, bright pupils appeared to be helped more by phonic instruction than those who were dull, and, although children with phonic training had some advantage in learning to spell, the non-phonic first-graders lost less during vacations.

Whilst experiments have shown that it is unprofitable to introduce all beginners to phonics, nevertheless it is reasonable to suppose that children of high intelligence and good cultural background might be expected to start earlier. More recent research on the optimum time for beginning phonic training, therefore, has largely been directed towards ascertaining the relationship between phonic ability and mental maturity. As yet, data on this topic are rather meagre, but such evidence as exists confirms that these two characteristics are highly correlated.

Dolch and Bloomster, for example, administered tests of phonic ability to first and second grade children in a school where phonics instruction had been given but not to an unusual degree. The results showed that, although able children sometimes fail to acquire phonic ability, those with a mental age below seven consistently fail to do so. When the experiment was repeated a year later, substantially the same results were obtained and so it was concluded that a mental age of seven years is necessary for a child to make the best use of phonics.

These findings are confirmed by Schonell and many other investigators, who likewise agree that certain developments and acquisitions are necessary before phonic training should be undertaken. According to Harris the important aspects of phonic readiness in increasing order of difficulty are as follows:

1. The child should be able to hear the difference between words that sound somewhat alike, such as *man* and *men*, or *had* and *hat*.
2. He should be able to detect whether two words begin with the same sound and supply two or three other words that begin with the same sound.
3. He should be sensitive to rhymes, able to pick out words that rhyme, and supply words to rhyme with a given word.
4. He should be able to hear similarities and differences in word endings and middle vowels of words.
5. He should be able to listen to the pronunciation of a word sound by sound, and fuse or blend the sounds mentally so as to be able to recognise the word intended.

Laboratory experiments have demonstrated that auditory discriminations, such as those listed above, are usually more difficult and therefore acquired at a later stage than those involving visual perception. Further, the visual perception of whole words has been shown to precede the ability to differentiate between the individual letters which form them. This additional evidence of the development of perceptual skills reinforces the conclusion reached by research in the classroom that *formal* phonic training for a considerable proportion of very young children is largely a waste of time.

The nature and amount of phonics instruction which should be given

Research studies have established the need for phonic readiness but it is nowhere suggested that teachers should not provide for its development in the reception class, or even earlier in the nursery school. Some researchers, including Dolch, Gray, Harris and Watts, have stated that at quite an early age *informal* training in the recognition of similarities and differences in word sounds by means of rhymes, etc., is profitable. It would seem that how phonic training is given is just as important as the timing of its introduction.

In the experiments previously cited Gates, Sexton and Herron concluded that it is preferable to teach sounds in words and not to drill them in isolation. Other investigators agree with this finding but seem to disagree about the amount of phonic instruction which should be given, and whether it should be formal or informal.

In 1938, Gates and Russell reported the results of a study in which 354 first-graders " fairly representative of the population of New York City," were first equated for mental age and then assembled into three groups. Group D " received the smallest amount of phonics or word analysis." Group E was given " moderate amounts of informal, newer type word analysis, comparisons, etc." Group F " had substantial or large amounts of conventional phonic drill."

At the end of the training period Group E had the highest scores in all four tests of word recognition and comprehension, Group D exceeded Group F slightly in two tests and was equal in the rest. It was concluded that a programme including little or no phonic analysis in the first year is not as good as an informal programme but is at least as good as one containing large amounts of formal phonic work.

Tate, Herbert and Zeman also found that the incidental teaching of phonics in connexion with children's needs in working out words encountered in their reading was superior to the teaching of isolated phonics or no phonics instruction.

When Agnew considered the effects of large amounts of phonic training on grade one children, he too found a slightly detrimental effect. In grade two the effect seemed to be one of slight reading improvement and in grade three of more accurate pronunciation. Since he was unable to establish a critical grade in which phonic experience is particularly effective, a further experiment was carried out. This time he reported that third-grade children who had experienced greater amounts of phonic instruction excelled, by statistically reliable differences, in tests of word pronunciation, vocabulary, accuracy of oral reading and eye-voice span, and yet there was comparatively little difference between them and pupils receiving a small amount of phonic instruction in the silent reading abilities measured.

The kind of reading skill developed by systematic phonic instruction

Important as it is to be able to recognise words in order to read them, it might be argued that comprehension is the essence of reading and that a lopsided development of word recognition skills has no merit, *per se*.

This argument is supported by the work of Tate who concluded that: "The use of as many as thirty minutes daily for special phonics instruction and drills leads to an unbalanced development of the abilities to comprehend words, to understand sentences and to grasp the meaning of paragraphs."

More recently, a similar conclusion was reached by McDowell in a study of fourth-graders of comparable intelligence and socio-economic background from two groups of Pittsburgh schools. He reported that in the group of schools where phonic training was introduced gradually, developed through the analysis of meaningful material and treated as a subsidiary word-attack skill, pupils were able to read faster, understand words and paragraphs, use an index, and generally read better than those in schools employing a conventional phonics programme. The phonic group was clearly superior in alphabetizing and spelling but, according to the experimenter, this superiority had been achieved at the expense of a more rounded and harmonious development of reading skills.

These researches and the work of other investigators, including that of Garrison and Heard and Gates mentioned earlier, would seem to conflict with Agnew's finding that improved word recognition through formal phonics training does not result in a lowering of reading comprehension ability.

There is, however, some evidence to support Agnew's conclusions in the recent investigation conducted in Kent primary schools by the National Foundation (Morris). Here, children aged seven to eleven who had attended schools which concentrated on systematic phonic instruction at the infant stage, were markedly superior in tests of reading *comprehension* to those who had been pupils at schools where phonic training had been given incidentally and without emphasis. These results were obtained under 'normal' school conditions (i.e. without recourse to 'artificial' experiment), and after allowance had been made for the children's non-verbal ability, and three school characteristics,

namely, urban/rural location, size of school, and the socio-economic status of the school catchment area.

As Daniels and Diack suggest, it seems likely that much of the controversy arises because different varieties of the phonic method have been used. It is also probable that sufficient attention has not been paid to differences between the immediate and after-effects of these methods, for Tiffin and McKinnis discovered that there is a high positive correlation between reading comprehension ability and phonic ability when children are in grades five to eight. They therefore concluded that a reading programme " which does not, by direct or indirect instruction, yield a mastery of the principles of phonics is not accomplishing its full purpose."

The type of pupil who benefits most from phonic teaching

Phonic ability is not necessarily the result of satisfactory school methods for, according to Schonell and others, intelligent children need comparatively little phonic teaching since they are able to understand and blend word sounds for themselves. It is children of poor intelligence, physique, emotional stability and socio-economic status for whom careful attention to methods is important.

It will be remembered that as early as 1913 Valentine showed that the dull child might make greater progress with look-and-say rather than phonic methods in the early stages. Ten years later Currier likewise pointed out that whilst phonic drills are valuable they are not essential to every child, and that at all times they should be employed with discretion and adapted to the needs of individual children or special groups.

Experiments designed to find out more about phonic readiness have indicated that formal phonic training should be delayed until a certain maturity has been achieved. This means that some children at seven plus, or even older, are only just ready to receive systematic phonic instruction, but that if this is not given, satisfactory reading progress may not be made. Some investigators have therefore turned their attention to the problem of phonic training in relation to children beyond the infant stage. In this connexion, Daniels and Diack compared the effects of teaching reading to first-year junior non-readers by the modified phonic method used in the *Royal Road Readers*, and a ' mixed-methods ' approach beginning with look-and-say methods supplemented at some stage by phonic teaching. After one year the children were given six specially designed tests and it was found that those taught by the phonic word-method were markedly superior in word recognition and word attack to pupils in the mixed-methods group.

The value of phonic teaching with backward readers in second year junior B and C classes has also been shown in an investigation by Conduct and Ward. Besides their ordinary reading of books the children were given apparatus, games, and tests of phonic sounds, and were also taught to recognise sounds in words and how to attack new words. After a ten-month period of teaching, which included individual, group and class activities, the reading ages of the 44 pupils were increased by 0.8-3.0 years, the average increase being 1.5 years. Furthermore, all achieved some degree of literacy although some of the children were almost non-readers when the experiment began and several of them were sub-normal in intelligence.

Not always has the effectiveness of phonic methods with backward readers been so clearly demonstrated. In fact, after experimenting with ten to eleven-year-olds drawn from their normal classroom environment, Burt and Lewis reported that, on the whole, the best method of teaching dull and backward readers is one which is predominantly visual. They pointed out, however, that the differences found in their experiment between whole-word, alphabetic, phonic, kinaesthetic and 'mixed' methods, even when significant, are still not very large, and that there is an urgent need for more intensive research into the problem.

It is possible, as Conduct and Ward suggest, that research into the effectiveness of different remedial measures has been directed too much to the individual and too little to the class. Individual treatment is known to give positive results whatever the methods used. What inexperienced teachers need to know is which methods are most effective in ordinary classroom conditions where it is often difficult to give sufficient attention to the needs of individual children. As will have been observed, some progress has been made towards answering this complex question in the limited field of phonics.

Summary

We now know that phonic instruction is of value and should be included in the reading programme. To be successful, however, it must be regarded as only one method of developing word skills and should not be introduced formally until children are ready for it. Bright pupils usually will have acquired sufficient vocabulary and reached sufficient maturity to be able to start phonic analysis in the reception class, but for those who are dull or handicapped in any other way, formal phonic training can profitably be postponed until a later stage in the school course—even into the first or second junior school year. There is still some doubt about the amount of phonic training which should be given but convincing evidence has been obtained that sounds should be taught in words and not drilled in isolation. Phonic methods undoubtedly improve word recognition but it is not certain whether this is gained at the expense of comprehension and interest in reading; the high positive correlation between phonic ability and reading comprehension ability in older children would suggest that this might not be so. Finally, the effectiveness of phonic methods with backward readers beyond the age of seven is indicated by the few experiments which have been carried out in conditions approximating to those of a normal class.

Bibliography

Agnew, D. C., "The Effects of Varied Amounts of Phonetic Training on Primary Reading." *Duke University Research Studies in Education*, No. 5. Durham, N.C.: Duke University Press, 1939.

Burt, C., and Lewis, R. B., "Teaching Backward Readers." *Brit. J. Ed. Psych.*, Vol. XVI, Part III, November, 1946.

Conduct, G. N., and Ward, H., "Phonic Reading—A New Approach." *Educational Review*, Vol. VII, No. 3, June, 1955.

Currier, L. B., "Phonics or No Phonics." *The Elementary School Journal*, Vol. XXIII, February, 1923.

Daniels, J. C., and Diack, H., *The Royal Road Readers*. London: Chatto & Windus, 1954.

Daniels, J. C., and Diack, H., *Progress in Reading*. University of Nottingham Institute of Education, 1956.

Dolch, E. W., *Problems in Reading*. Champaign, Illinois: Garrard Press, 1948.

Dolch, E. W., and Bloomster, M., "Phonic Readiness." *The Elementary School Journal*, Vol. XXXVIII, November, 1937.

FERNALD, G. M., and KELLER, H. B., " The Effect of Kinaesthetic Factors in the Development of Word Recognition in the Case of Non-Readers." *J. Educ. Res.*, Vol. IV, 1921.
FLESCH, R., *Why Johnny Can't Read.* New York: Harper, 1955.
GARRISON, S. C., and HEARD, N. T., " An Experimental Study of the Value of Phonetics." *Peabody Journal of Education*, Vol. IX, July, 1931.
GATES, A. I., " Studies of Phonetic Training in Beginning Reading." *J. Ed. Psych.*, Vol. XVIII, April, 1927.
GATES, A. I., and RUSSELL, D. H., " Types of Material, Vocabulary Burden, Word Analysis and Other Factors in Beginning Reading." *The Elementary School Journal*, Vol. XXXIX, 1938.
GILL, E. J., " Methods of Teaching Reading: A Comparison of Results." *J. Exp. Pedagogy*, Vol. I, No. 3, March, 1912.
GRAY, W. S., " The Nature and Organisation of Basic Instruction in Reading." *The Teaching of Reading: A Second Report*, 36th Yearbook of the U.S.A. National Society for the Study of Education, Part I. Bloomington, Illinois: Public School Publishing Co., 1937.
GRAY, W. S., " Reading." *Encyclopedia of Educational Research*, 1950.
HARRIS, A. J., *How to Increase Reading Ability.* New York: Longmans Green, 1947.
MCDOWELL, J. B., " A Report on the Phonetic Method of Teaching Children to Read." *The Catholic Educational Review*, Vol. LI, October, 1953.
MORRIS, J. M., *Reading in the Primary School.* Publication No. 12 of the National Foundation for Educational Research. London: Newnes (*in the press*).
MOSHER, R. M., and NEWHALL, S. M., " Phonic Versus Look-and-Say Training in Beginning Reading." *J. Ed. Psych.*, Vol. XXI, October, 1930.
RUSSELL, D. H., " Teachers' Views on Phonics." *Elementary English*, Vol. XXXII, October, 1955.
SCHONELL, F. J., *Backwardness in the Basic Subjects.* Edinburgh: Oliver & Boyd, 1942.
SCHONELL, F. J., *The Psychology and Teaching of Reading.* Edinburgh: Oliver & Boyd, 1947.
SEXTON, E. K., and HERRON, J. S., " The Newark Phonics Experiment." *The Elementary School Journal*, Vol. XXVIII, May, 1928.
TATE, H., " The Influence of Phonics on Silent Reading in Grade I." *The Elementary School Journal*, Vol. XXXVII, June, 1937.
TATE, H., HERBERT, T., and ZEMAN, J., " Non-Phonic Primary Reading." *The Elementary School Journal*, Vol. XL, March, 1940.
TIFFIN, J., and MCKINNIS, M., " Phonic Ability: Its Measurement and Relation to Reading Ability." *School and Society*, Vol. LI, February, 1940.
TRAXLER, A. E., *Ten Years of Research in Reading.* Educational Records Bulletin, No. 32. New York: Educational Records Bureau, 1941.
TRAXLER, A. E., and Townsend, A., *Another Five Years of Research in Reading.* Educational Records Bulletin, No. 46. New York: Educational Records Bureau, 1946.
TRAXLER, A. E., and TOWNSEND, A., *Eight More Years of Research in Reading.* Educational Records Bulletin, No. 64. New York: Educational Records Bureau, 1955.
VALENTINE, C. W., " Experiments on the Methods of Teaching Reading." *J. Exp. Pedagogy*, Vol. II, No. 2, June, 1913.
VERNON, M. D., *Backwardness in Reading.* Cambridge University Press, 1957.
WATTS, A. F., *The Language and Mental Development of Children.* London: Harrap, 1944.
WINCH, W. H., " Teaching Beginners to Read in England: Its Methods, Results and Psychological Bases." *Journal of Educational Research Monographs*, No. 8. Bloomington, Illinois: Public School Publishing Company, 1925.

A COMPARATIVE STUDY OF BASIC ATTAINMENTS

The Results of Surveys of the Attainments in Reading and Arithmetic of Eleven-year-old Children in England and Wales; Queensland, Australia; and Central California, U.S.A.

by D. A. PIDGEON

I. Introduction

IN May, 1955, as part of a national survey involving three age groups, a representative sample of some 3,000 children, aged eleven plus, attending primary schools in England and Wales, were given specially prepared tests of reading, arithmetic and non-verbal ability. After reading an account of the proposed survey in the Foundation's Bulletin No. 5, the Principal Research and Guidance Officer for the State of Queensland, Australia, sought permission to apply the same tests to a comparable sample of eleven-year-old children in Queensland. The request was granted and in November, 1955, the survey was repeated on a representative sample of children from that State. In the summer of 1956, following a visit to the Foundation by Professor G. T. Buswell of the University of California (Berkeley), U.S.A., arrangements were made for one of the arithmetic tests used in the Foundation's survey to be repeated in Central California. Accordingly, in January, 1957, this arithmetic test was given to a representative sample of eleven-year-old Californian children.

In the following pages an account is given of a comparision of the results obtained from these three surveys.

II. The Samples Tested

(*a*) ENGLAND AND WALES

The national survey was confined to grant-aided schools maintained or assisted by local education authorities in England and Wales, and in this particular age group was restricted to children attending junior and all-age schools who had reached their tenth birthday but not their eleventh by August 31st, 1954, that is, who were born between September 1st, 1943, and August 31st, 1944 inclusive. Since the testing was carried out in May, 1955, the effective age range of the children was from 10 years 9 months to 11 years 8 months. A random sample of schools was drawn and all children in the age group in each selected school were tested. The selection of schools was made directly from the lists kept at the Ministry of Education of all schools in England and Wales attended by the children concerned. In order to ensure the adequate representation of schools of different types and in urban and rural areas, random sampling was carried out separately in each of these strata. The sampling method was thus systematic but unbiassed and was considered to be effectively random. In all, some 3,180 children attending 91 schools administered by 63 different local education authorities were tested.[1]

[1]Full details of the sampling procedure employed will be published later.

(*b*) Queensland, Australia

Surveys of children's abilities and attainments had previously been carried out in Queensland and therefore the technique of obtaining effectively random samples was not new to them. In this inquiry, as in England and Wales, the schools selected for the survey were a stratified random sample. The testing was confined to children born between March 1st, 1944, and February 28th, 1945, and to equate conditions with regard to the school year in this country the testing was carried out in November, 1955. Thus, the mean age of the sample was 11 years 2½ months and the fact that the tests were given towards the end of the school year made it virtually identical with that used in England and Wales. The schools in Queensland containing children in the appropriate age group were divided into the five groups according to size that are used officially for purposes of classification. The groups are: Class I, average attendance over 400; Class II, 150-400; Class III, 41-150; Class IV, 21-40; Class V, 20 and under. The number of schools selected in each class was fixed at 1/17th of the number of schools in that class in 1954. As this would have led to the selection of only six schools of more than 400 pupils, it was decided to select three times as many Class I schools but to test only one-third of the children in each. The grade system of placement is used in Queensland schools and although approximately two-thirds of the children in the defined age group were in Grade V the others were found as low as Grade II (two pupils) and as high as Grade VII (one pupil). In all some 940 children attending 88 schools were tested.

(*c*) California, U.S.A.

The sampling procedure in California was very similar to that employed in England and Wales. Again, although the grade system of placement is used in America, the testing was confined to children of the same age range at the date of testing as those in this country and in Queensland. Thus, the sample was drawn from all pupils who were born between May 1st, 1945, and April 30th, 1946, inclusive, and the testing was carried out in January, 1957. The 773 urban and 661 rural schools in the 21 counties in Central California were numbered serially, and after stratifying according to urban or rural locality an effectively random sample of schools was selected to give a sample of children approximately equal in number to that obtained in England and Wales. Invitations were issued to the selected schools to participate in the experiment and only three of the school systems refused. One or two other minor exceptions to the sampling procedure were necessary, but it is unlikely that they exerted any appreciable effect upon the results obtained. In all, 3,179 pupils from 70 schools were tested.

III. The Tests

Four tests were administered in the Foundation's national survey to children in the 11+ age group. These were:

1. *A test of non-verbal ability*, consisting of 64 items of two main types. The first type consisted of a large square with the lower right-hand corner missing, and several smaller squares from which the appropriate one had to be selected to complete the large square. The second consisted of a sequence of three or four drawings with the next one missing, and a group of five more from which the appropriate one had to be chosen to continue

the sequence. The items were arranged four to a page, but they were not grouped in separate series. A practice test of eight items was administered with detailed oral instructions before the test proper. The reliability of the test was found to be 0.943.*

2. *A reading comprehension test*, consisting of 60 incomplete sentences. The appropriate word necessary to complete the sentence had to be chosen from five alternatives. The items, presented in order of difficulty, ranged from sentences suitable for bright seven-year-olds to those found difficult by children of fourteen. The test was scored by counting all correct items up to the point at which a gap of seven consecutive wrong answers occurred. Any items correct beyond this point were assumed to be due to chance and ignored. The reliability of this test was 0.958.*

3. *A mechanical arithmetic test*, consisting of twenty-five items ranging from a two row addition sum with hundreds, tens and units, and a carrying figure in the units column, to a simple multiplication of decimals. Ten sums dealt with number facts only, the most difficult being a long multiplication and a long division without remainder. A further nine sums dealt with conversion units, involving money, time, length, weight and capacity. The last six items consisted of four elementary fraction sums, one in each basic process, and two simple decimal sums. The test was administered without time limit. The reliability of this test was 0.922.*

4. *A mechanical and problem arithmetic test* consisting of 100 items administered with a time limit of 50 minutes. The test was in two sections, separately timed. Section 1 consisted of 40 mechanical sums involving the four rules and exercises in money, weight, time, length and capacity. Section 2 consisted of 60 simple one-step problems covering the same ground with the addition of one or two involving elementary fractions and area. The reliability of this test was .982.*

These four tests were also given in the same order and with the same instructions in Queensland, Australia.

In California only the second 100 item arithmetic test was administered. Certain changes were necessary in adapting the test for American usage. (*a*) Seventy of the 100 items were considered suitable for use in American schools with only minor verbal changes. (*b*) Twenty-six items involved money terms. These were changed to dollars and cents, with due regard to the size of the numbers and the arithmetical processes involved. (*c*) The remaining four items involved weight measured in "stones"—a unit not used in America. In tabulating the results, both in California and in this country, the scores obtained in these three groups of items were treated separately. For this present report an analysis was carried out on the 70 comparable items only, of which 31 were mechanical sums in Section 1, and 39 were problems from Section 2. On the data from England and Wales, the reliability was .974 for the total of 70 items, and .949 for Section I and .955 for Section 2 separately. On the Californian data the reliability was .89 for the total of 70 items and .68 for Section 1 and .87 for Section 2 separately.*

*By the reliability of a test is meant the degree to which scores derived from it are likely to be consistent. Measures of reliability are usually expressed in a scale from 0 to 1.0. The measures for these tests were obtained by Kuder-Richardson formula 20.

IV. Administration of the tests

(*a*) England and Wales

In the Foundation's national survey the testing was carried out, either by the head teacher of the school concerned, or by a psychologist or other trained official of the local education authority concerned. In the case of the former, responsibility for the administration of the tests rested nevertheless with an official of the local education authority. A detailed instruction booklet with clear and simple directions for giving the tests had been prepared. The tests were administered on the mornings only of three consecutive days; the non-verbal test was given on the first day, the reading test, followed after a short break by the untimed arithmetic test, was given on the second day, and the timed arithmetic test was given on the third day.

In order to ensure that the testing was properly carried out each local authority concerned was asked to appoint an official to be responsible for the conduct of the survey locally. These officials were briefed by members of the Foundation at a series of regional meetings held for the purpose at eight centres scattered over the country. Reports on the conduct of the survey locally were asked for and a study of these indicated that in most cases a great deal of trouble had been taken to ensure that the children worked in circumstances that enabled them to do their best. Although in particular instances, especially in small rural schools, difficulties of accommodation were experienced, it appeared that these were overcome, and reliance could be placed upon all scores obtained. Completed tests were returned to the Foundation for marking and checking.

(*b*) Queensland, Australia

In Queensland the order of giving the tests and the instructions for administration were identical with those in England and Wales, and considerable care was taken to ensure that conditions were satisfactory. The tests were administered by teachers, but the marking and checking was carried out by the staff of the Research and Guidance Branch of the Department of Public Instruction.

(*c*) California, U.S.A.

In California all the tests were administered by Professor Buswell or his research assistant, George Yonge. Generally, one or more teachers or supervisors were also in the room during the testing and the physical conditions were reported to be good. The instructions for the test were given orally and were uniform with those used in England. In some of the schools the test was given in the regular classrooms, pupils outside the age range being sent to other parts of the building, while in the larger schools the pupils to be tested were assembled in a multi-purpose room using cafeteria or library tables. All the tests were scored by clerical assistants in the research office at Berkeley University. Their work was checked systematically to ensure correct scoring.

V. Results and Comments

(*a*) Comparison between England and Wales and Queensland

As only one test was given in California, the results of the full survey common to both this country and Queensland will be given first.

The basic sample of children considered in each country was that consisting of those children who completed all four tests. In England and Wales, the data obtained from all children who completed each test were, however, used to standardise the tests. The method used for the standardisation was an adaptation of that devised by Lawley[1] in which the standardised scores are so arranged that they have a mean of 100 and a standard deviation of 15.[2] Using the conversion tables prepared, the mean standardised scores of the sample of children taking all tests were thus calculated for both England and Wales and Queensland. Table 1 gives these means (standard errors in brackets), together with the respective standard deviations.

Table 1

COMPARISON OF MEAN STANDARDISED SCORES FOR FOUR TESTS GIVEN IN QUEENSLAND, AUSTRALIA AND IN ENGLAND AND WALES

	N/V Test		Reading Test		Untimed Arithmetic Test 1		Timed Arithmetic Test 2	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Queensland, Australia	94.7 (0.66)	12.5	95.3 (0.65)	12.7	100.3 (0.54)	10.6	99.7 (0.57)	10.4
England and Wales	100.48 (0.80)	14.50	100.03 (0.69)	14.51	100.32 (0.80)	14.88	100.39 (0.77)	14.75

The slight discrepancies between the figures given for England and Wales in Table 1 and the rounded values of 100 and 15 for the means and standard deviations respectively are due in the main to the omission of partial absentees. The mean standardised scores of those children who were present for one test, but absent for one or more of the others, were in fact calculated and found to be some 4 or 5 points below the value of 100.

The main fact emerging from Table 1 is the significantly lower mean scores obtained by the Queensland children on both the non-verbal and reading tests. The differences between the means for both arithmetic tests are not significant. The appreciably smaller standard deviations are also worthy of note.

In any discussion of these findings two comments received from Queensland are relevant. In the first place it was pointed out that this sample of Australian children were completely lacking in test sophistication and that this fact might affect the scores on the non-verbal test more than those on the other tests. Secondly, previous Australian comparisons had shown Queensland pupils to be well ahead of those in the other States on standardised arithmetic tests (of a type similar to the untimed Arithmetic Test 1) and average on tests of silent reading and non-verbal intelligence.

It may well be true that the English sample of children had enjoyed a wider experience of tests in general, although from the information gleaned from the survey reports sent

[1] *A Method of Standardising Group Tests* by D. N. Lawley, B.J.P. (Statistical Section), Vol. III, Part II, June 1950.
[2] The 'standard deviation' is a measure of the spread of the scores—i.e. the extent to which they tend to deviate from the mean or average

in by many local authorities it is unlikely that many of the children tested had previously been given non-verbal tests or tests of reading similar to the one employed in the survey.

The non-verbal test

It is possible, however, that the children in England and Wales were more familiar with the type of perceptual material employed in the non-verbal test by reason of the increasing number of puzzle toys now available for children of primary school age. In any case, it is important to remember that this test is measuring only one aspect of intelligent behaviour, and the different cultural environments of the children in the two countries may well be responsible for the difference in the scores obtained. It cannot be assumed, therefore, that the lower scores of Queensland children in any way reflect a lower innate intelligence.

It will be observed in Table 1 that the standard deviations of the standardised scores for the non-verbal and the other three tests are somewhat smaller in the Queensland data than they are in England and Wales. A more detailed comparison of the distribution of test scores in the two countries was made and Table 2 gives the proportions found in three groups of standardised score. The expected proportions from a normal distribution are also given.

Table 2

PROPORTIONS IN QUEENSLAND AND IN ENGLAND AND WALES OCCURING IN THREE GROUPS OF STANDARDISED SCORE, ON FOUR TESTS

Range of Standardised Scores	Expected Proportion from Normal Distribution	Non-Verbal Test		Reading Test		Arithmetic Test 1		Arithmetic Test 2	
		Q.	E.	Q.	E.	Q.	E.	Q.	E.
115 & above	15.87	5.73	17.20	5.73	14.89	8.38	17.14	6.04	16.54
85–114	66.52	76.88	68.66	75.40	69.63	85.47	68.47	85.47	68.17
84 & below	17.62	17.40	14.13	18.87	15.48	6.15	14.40	8.48	15.28

It is clear from Table 2 that the lower mean score on the non-verbal test for Queensland children was due, in the main, to the relatively smaller proportion of high scorers.

Reading Test

So far as the reading test is concerned the possibility that some of the words used were inappropriate for Australian children cannot be overlooked. It was ascertained from the Principal Research and Guidance Officer, Queensland, however, that although an occasional word or concept might be more fitting for English than Australian children, he would hesitate to assert this. An item analysis was carried out on a sub-sample of scripts from each of the two countries, and an inspection of the results showed that in only one item was there a serious discrepancy. This item, answered correctly by approximately fifty per cent. of English children and by only twenty-five per cent. of Queensland

children, was concerned with the meaning of the word " conjuror " which presumably is not so familiar in Australia.

The proportions of high and low scorers on this test, given in Table 2, are interesting. Not only is there a much higher percentage of English and Welsh children who do well on this test, but there is a slightly smaller percentage doing badly. It may be safe, therefore, to conclude that in-so-far as this test measured reading comprehension, the children from this country were superior in this respect to those from Queensland.

Arithmetic Tests

The values for the mean scores given in Table 1 on both arithmetic tests would suggest that there is little to choose between the performances of children from the two samples, although the low standard deviations of the Queensland children call for attention. The appropriate figures in Table 2 emphasise the difference between the distribution of scores obtained in the two countries. There are more high scorers on both tests among the English and Welsh children, but there are also more low scorers. In seeking an explanation of these findings it would seem appropriate to examine the different curricula followed by the two groups of children in their respective primary schools. Before doing so, however, it is necessary to look at another factor that might have influenced the results.

If it can be assumed that the sample children from this country were relatively unsophisticated in non-verbal and reading comprehension tests, a similar assumption could not be made with regard to the timed Arithmetic Test 2. This test was similar to that employed by many local education authorities as part of their 11 + allocation procedures, and since nearly all the English children had undergone a selection examination some three months previously, this type of test could not be regarded as unfamiliar. Not all authorities, however, employed standardised arithmetic tests in their allocation procedures, but from the information available at the Foundation it was possible to divide the sample schools into two groups: Group (*a*)—those in authorities that regularly used a standardised arithmetic test, and Group (*b*)—those in authorities that did not. The mean standardised scores, on all four tests used in the survey, of the children attending schools in these two groups are given in Table 3.

Table 3

MEAN STANDARDISED SCORES ON FOUR TESTS FOR:

(*a*) Children in L.E.As. that used standardised arithmetic tests.

(*b*) Children in L.E.As. that did not use standardised arithmetic tests.

	Non-Verbal Test	Reading Test	Untimed Arithmetic Test 1	Timed Arithmetic Test 2
Group (*a*)	101.03	100.18	100.67	101.14
Group (*b*)	99.85	99.86	99.93	99.52
Difference	1.18	0.32	0.74	1.62

None of the differences shown in Table 3 is significant, although, as might have been expected, the largest difference is found with the timed Arithmetic Test 2.

The influence of test sophistication may, however, be seen in another way. Since Arithmetic Test 2 is timed and the rate of working expected is rather high, it is possible that English children as a whole may have been able to tackle more of the test in the time allowed than their counterparts in Queensland. This point was investigated by comparing the average facility values[1] for consecutive groups of ten items obtained from item analyses carried out on sub-samples of the completed scripts. On the English and Welsh sample, a one-in-seven random sample of the original scripts was used for this purpose, and in Queensland the indices of difficulty were obtained from a sample of 150 pupils whose scores on the non-verbal test were representative of those of all pupils.[2] Table 4 gives the average facility value for successive groups of 10 items.

Table 4

MEAN FACILITY VALUE FOR SUCCESSIVE GROUPS OF 10 ITEMS IN ARITHMETIC TEST 2, FOR QUEENSLAND AND ENGLAND AND WALES

Items	Mean Facility Values	
	Queensland	England & Wales
Section 1. Items 1–10 ..	76.2%	69.2%
,, 11–20 ..	64.0%	52.3%
,, 21–30 ..	48.6%	38.6%
,, 31–40 ..	16.0%	15.9%
Average for Section 1 ..	51.2%	44.0%
Section 2. Items 1–10 ..	67.6%	69.6%
,, 11–20 ..	42.2%	44.9%
,, 21–30 ..	46.2%	42.6%
,, 31–40 ..	37.6%	37.8%
,, 41–50 ..	15.6%	23.1%
,, 51–60 ..	4.6%	13.3%
Average for Section 2 ..	35.6%	38.6%
Test Average	41.9%	40.7%

There are a number of comments to be made on Table 4. In the first place the mean raw score on this test was 40.5 in both countries. It is clear, therefore, that the Australian sub-sample used in this analysis was slightly superior to the total sample of Queensland children. However, this does not affect the other comments to be made, namely, that

[1] The 'facility value' of an item is denoted by the percentage of children in a sample who answer the item correctly.

[2] As would be expected in this case, the distribution of scores on Arithmetic Test 2 showed a slight regression to the mean, but it was not thought that this would affect the indices to any great extent.

the Australian children did considerably better than their English and Welsh counterparts on Section 1—the mechanical section—of the test, were only about equal to them on Section 2—the problem section—and failed, on the average, to get as many items completed towards the end of the test. The assumption, necessary for this last comment, appears to be justified, since there are no specifically new types of item introduced towards the end of the test and the actual increase in difficulty is only slight. It appears, therefore, that, although the mean standardised scores for the two countries on this arithmetic test did not differ significantly, this was due (*a*) to the superior performance of the Australian children on mechanical sums, and (*b*) to the English and Welsh children being able to get more of the problem sums right towards the end of the test.

However, the question of rate of working does not enter into the results for Arithmetic Test 1, since it was administered without time limits. Also, since this test consisted of mechanical sums only, a better all-round performance on the test might have been expected from Queensland children. But this is not so, and in order to discover the reason why, it is necessary to examine in greater detail the respective performances of the children from the two countries on specific types of item. Arithmetic Test 1 was, in fact, included in the survey for this very purpose. It is not possible in giving the results of the item analysis to quote the actual sums used in the test, since it was constructed specially for the survey and it is hoped to repeat it on a later occasion. Neither is it considered necessary to detail all 25 items. Table 5, however, gives the comparative proportions answering correctly for 10 selected items. The Australian data are based on the sub-sample of 150 children described previously; the values for the English and Welsh children were derived from all the 3,185 children taking the test.

Table 5

PERCENTAGES OF CHILDREN IN QUEENSLAND AND IN ENGLAND AND WALES ANSWERING CORRECTLY SELECTED ITEMS FROM ARITHMETIC TEST 1

Item No.	Type of Item	Facility Values	
		Queensland %	England %
5	Simple add. 3 row h.t.u. 2 carrying figs.	96	92
6	Simple sub. h.t.u. borrowing in 1 col. and zero ..	83	79
7	Simple mult. h.t.u. by u.	88	77*
8	Short div. th.h.t.u. by t.u. no remainder	89	74*
13	3 row add. hr. min. sec.	66	59
14	Sub. yd. ft. in. borrowing in 1 col.	68	68
15	Simple mult. gal. qt. pt.	55	47
16	Short div. tons cwt. lb. no remainder..	61	36*
21	Sub. simple fractions	21	36*
24	Add. of dec. 3 simple numbers presented horizontally	23	43*

Not all the differences shown in Table 5 are significant, but it seems apparent that the Australian children are generally better on all items except those involving fractions

and decimals. It will be recalled that there are three main types of item in this test: (*a*) those dealing with simple numbers, involving the 4 rules, including one long multiplication and one long division (10 items); (*b*) those dealing with conversion units, including money, time, length, weight and capacity (9 items): (*c*) those dealing with simple fractions (4) and decimals (2) (6 items). The mean facility values for these three types of item are given in Table 6.

Table 6

COMPARISON OF THE MEAN FACILITY VALUES FOR THREE TYPES OF ITEM BETWEEN CHILDREN IN QUEENSLAND AND IN ENGLAND AND WALES

No. of Items	Type of Item	Mean Facility Values	
		Queensland	England & Wales
10	Simple number: 4 rules	87.1%	79.0%
9	Conversion units. Money, time, length, weight and capacity	65.4%	58.1%
6	Simple fractions and decimals	17.7%	34.3%

The pattern of responses to items of mechanical computation is now quite clear. The children from Queensland give a superior performance to English and Welsh children on both arithmetic tests in all mechanical items except those involving simple fractions and decimals. The latter are not introduced in the schools in Queensland until the Sixth Grade and thus only about eighteen per cent. of the children in the sample had had any formal instruction in sums of this type. Had questions of this kind not been included, therefore, the Australian children would have obtained a higher mean score on Arithmetic Test 1 than the children from this country.

This finding is of interest—not because the English and Welsh children do better than Queensland children on items involving fractions and decimals, but because their success on this type of item is as high as it is. The fact that about one-third of the children from England and Wales were able to answer the fraction and decimal sums correctly is surprising, since this type of item does not appear in the mechanical section of the standardised tests of arithmetic used by the majority of authorities in their 11 + allocation procedures (cf. Arithmetic Test 2). It might have been expected, by reason of the alleged " backwash " effect of such tests, that exercises of this kind would not be taught in English primary schools, but the evidence from the survey is clearly to the contrary.

Teaching mechanical computation is not, of course, the sole constituent of the arithmetic curriculum in primary schools and the necessity for pupils to be able to answer written problems is equally important. Reference to the mean facility values (given in Table 4) for the first forty items of the second (problem) section of Arithmetic Test 2 would indicate little difference in the performances of children from Queensland and England and Wales. The variation in the type of arithmetical problem given in this test is, however, quite large and the grouping adopted for Table 4 does not give any

indication of possible differences that may occur between different types of problem. It would be desirable, therefore, to compare the average facility values for items of a similar kind, but in practice this is very difficult to achieve since the majority of problems can be classified under a number of different headings.

An inspection of the separate facility values, however, revealed interesting differences, and to illustrate these the first forty items of the second section of the test were divided into two main groups. Three items were omitted from this analysis since they involved fractions and, for the reasons stated above, a comparison involving those items would clearly have been inappropriate. The last twenty items were also omitted since it was clear that in these the Queensland children were handicapped by the imposition of strict time limits. The two groups, the mean facility values of which are given in Table 7, were as follows: Group (*a*) Eighteen simple written exercises involving the four rules and conversion units, e.g., item 2, " John has 7 sweets, and I have three times as many. How many have I? "[1] Item 6, " Write in figures the number eleven thousand two hundred and seven." Item 38, " Take 144 pence from 47 shillings and give your answer in £ s. d." Group (*b*) comprised nineteen more complex verbal problems, e.g., Item 4, " I bought a three shilling book of stamps and four 2½d. stamps. What was the total cost? " Item 16, " A man cycles at 12 miles an hour. How long will he take to travel 84 miles at this rate? " Item 39, " I can pay 8d. for a return ticket every day or buy a weekly ticket costing 4/-. How much shall I *lose* each week by buying a weekly ticket if I go to work only 5 days a week?"

Table 7

COMPARISON BETWEEN THE MEAN FACILITY VALUES FOR TWO GROUPS OF ITEMS FROM ARITHMETIC TEST 2 FOR QUEENSLAND AND ENGLAND AND WALES

	No. of Items	Type of Item	Mean Facility Values	
			Queensland	England & Wales
Group (*a*)	18	Simple written exercises in 4 rules and conversion units	58.9%	51.1%
Group (*b*)	19	More complex verbal problems ..	41.1%	46.1%

Of the eighteen items in Group (*a*) fourteen were answered correctly by more of the Australian children, and four by more of the children from England and Wales. Of the nineteen items from Group (*b*) the English and Welsh children did better in fifteen, the Queensland children in only one, and in three the facility values were the same.

In interpreting these findings it is necessary to remember that significantly lower scores were obtained by the Australian children in the reading comprehension test. Thus it is possible that in answering the more complex verbal problems, the Queensland children were handicapped to some extent by difficulties in reading comprehension. Alternatively, it is possible that the children from this country, because they had been prepared for tests of this kind, were more practised in answering verbal problems of this

[1] The numbers in the items quoted here have been changed slightly.

type. Whatever the explanation may be, the results obtained from the two arithmetic tests indicate that, despite the near equivalence of the mean scores, the Australian children are better able to perform many of the basic operations of arithmetic than children of the same age in England and Wales.

It is now possible to make a brief comment on the reasons for the markedly different distributions of test scores obtained by the Queensland children. It is clear that the relatively small proportion of children from Queensland scoring outside the approximate range of $\pm$ one standard deviation from the mean on the arithmetic tests (Table 2) is due in part to the content of these tests. They were made for children in this country and therefore included a number of items that were inappropriate for the Australian children. Nevertheless, it may be argued, that had these inappropriate items (fractions and decimals in the mechanical paper, and complex verbal problems in Arithmetic Test 2) been omitted, although the proportion of Queensland children scoring above 114 would have increased, the proportion scoring below 85 might well have decreased, giving further evidence of the superior performance in arithmetic of the Australian children.

Had the contents of these two arithmetic tests been equally appropriate for children from both countries, it might have been argued that the relatively greater dispersion of scores obtained in England and Wales was due to the system of streaming that is fairly widely employed—because of the demands of the 11 + examination "A" stream children are often given the best teachers and are encouraged to cover the syllabus at a rapid rate, whilst " C " stream children tend to be given less experienced teachers and have little incentive to progress. The detailed analysis of the arithmetic tests, however, suggests that this is not the complete explanation, and this conclusion is supported by the results on the reading test. Although the proportion of " good readers " (above 114) in this country is greater, suggesting better teaching for the abler pupils, the proportion of poor readers (below 85) is slightly less than in Queensland.

It appears, therefore, that the results of the reading and arithmetic tests taken together offer no conclusive evidence for the superiority of either system of organisation. The relatively larger number of " good readers " and " good arithmeticians " in this country may in part be due to the greater pressure brought to bear on "A" stream children of this age because of the 11 + examination. At the other end of the scale, however, there are inconsistencies between the subjects tested. Arithmetic appears to be better taught in Australia than in England and Wales in that there are fewer low scorers on the tests, but the same cannot be said for reading. The differences in performances on the attainment tests may well be due, therefore, to the methods and approach employed in teaching the different subjects and not to any overall differences in the organisation of classes.

There still remains one fact worthy of comment. Although it might be possible to explain the observed differences in the attainments of the children from the two countries in terms of differences in the effectiveness of the instruction they had received, such an explanation could scarcely be valid with regard to the results of the non-verbal test of ability. Nevertheless, there is evidence, from other sources, to show that performance in tests of this kind is affected by cultural influences, and it is reasonable, therefore, to suppose that the differences found in the results of this test, as well as those in the attainment tests, may be ascribed to environmental factors.

(*b*) Comparison between England and Wales and California[1]

Only one test is being considered here—the timed arithmetic Test 2—and since only a proportion of the items in this test could be regarded as being comparable it has not been possible to use standardised scores in the analysis. The results which follow, therefore, are given in terms of raw scores. It will be recalled that the number of strictly comparable items was 70, and the number of children contributing to the results was 3,191 from 91 schools in England and 3,179 from 70 schools in Central California. Table 8 gives the means and standard deviations of the scores from the two countries.

Table 8

MEANS AND STANDARD DEVIATIONS OF COMPARABLE ITEMS IN AN ARITHMETIC TEST GIVEN IN CENTRAL CALIFORNIA AND IN ENGLAND AND WALES

	California		England & Wales		Difference of Means
	Mean	S.D.	Mean	S.D.	
Section 1 Computation (31 items)	4.2	2.1	14.1	8.7	9.9
Section 2 Problems (39 items)	7.8	5.1	15.0	10.5	7.2
Total Test (70 items)	12.1	6.8	29.1	18.7	17.0

All the differences in Tab e 8 are highly significant.

As will be seen from Table 8 the differences between the two groups of children are very marked. A comparison of the distributions of total score emphasises the differences; the scores for the English and Welsh children spread over the entire range of the test from zero to a perfect score of 70, whereas there were no scores above 53 for the pupils from California. Thirteen pupils in California made scores above 38 as compared with 1,077 pupils in England who scored above this point.

From information obtained from California it was clear that this test, constructed as it was for children in this country, laid more emphasis in Section 1 on denominate numbers than is usual in American tests. An item analysis was therefore carried out which shows the percentage answering each of the 70 items correctly. The results are given in Table 9 and in order to complete the comparison the relevant figures for Queensland, Australia have been included with those from California and England and Wales. The item analyses were based on a random selection of 456 test papers from the English and Welsh sample and 544 from the Californian sample. The Queensland figures were obtained from the sub-sample of 150 children previously described.

[1] An account of this part of the investigation has been published in America. See "A Comparison of Achievement in Arithmetic in England and Central California," by G. T. Buswell, *The Arithmetic Teacher*, February 1958.

Table 9

PERCENTAGES OF CHILDREN FROM CALIFORNIA, QUEENSLAND AND ENGLAND AND WALES ANSWERING CORRECTLY EACH OF SEVENTY COMPARABLE ITEMS FROM ARITHMETIC TEST 2

a=California. *b*=Queensland. *c*=England and Wales.

Section 1 (Mechanical)								Section 2 (Problems)							
Item No.*	Per Cent. Correct			Item No.*	Per Cent. Correct			Item No.*	Per Cent. Correct			Item No.*	Per Cent. Correct		
	a	*b*	*c*		*a*	*b*	*c*		*a*	*b*	*c*		*a*	*b*	*c*
	%	%	%		%	%	%		%	%	%		%	%	%
1	85	91	93	19	22	61	43	1	89	81	91	32	7	43	47
2	80	87	84	22	0	58	45	2	85	91	84	33	9	41	45
3	67	80	76	23	1	59	44	3	65	87	73	34	7	33	34
4	59	91	78	25	3	48	42	5	66	37	59	36	19	30	33
5	7	65	61	27	0	46	37	6	55	87	69	37	17	40	31
6	66	83	70	28	0	39	31	9	23	48	54	41	4	29	29
7	4	61	59	30	0	33	24	12	42	61	46	42	7	17	30
8	0	69	58	31	0	35	26	13	15	27	61	44	7	24	21
9	2	70	64	32	0	28	22	15	13	68	37	45	3	17	24
10	3	65	49	33	0	21	22	16	24	40	48	46	5	21	23
11	2	57	48	34	1	16	17	17	19	39	51	47	3	13	22
12	1	71	57	35	0	19	18	18	4	23	34	48	7	9	25
14	4	65	62	36	0	15	15	20	21	43	43	49	1	4	12
16	0	60	49	37	0	10	14	21	7	58	40	52	5	7	12
17	2	71	57	39	4	5	9	22	33	32	47	55	1	3	11
18	1	63	47					23	39	73	52	56	1	9	21
								24	28	44	44	57	1	5	13
								25	7	53	45	59	1	2	9
								29	7	43	47	60	1	1	3
								30	6	35	34				

*Item numbers correspond to the numbers on the printed test form.

Table 9 shows that the percentage of English and Welsh children answering the item correctly exceeded that for the American children on all but two items, (Nos. 2 and 5 of Section 2). There were five items in the test which involved hundredweight. Although this term is included in the tables of the Californian State textbook there was no indication that it had been taught, and furthermore it is seldom used in America. The average per cent. correct on these five items, (Nos. 22, 27, 35 and 36 in Section 1 and No. 21 in Section 2) is 31 per cent. in England and Wales and 1.4 per cent. in California. If these five items were omitted from the test it would, however, reduce the mean score on the total test for English and Welsh children by approximately 1.6 points but would make no significant difference in the comparison.

Table 10 below gives some examples of the items employed in the test with the percentages from the three groups of children answering them correctly.

Table 10

SAMPLE ITEMS FROM ARITHMETIC TEST 2 WITH PERCENTAGES OF CHILDREN FROM CALIFORNIA, QUEENSLAND AND ENGLAND AND WALES ANSWERING CORRECTLY

Item No.	Item*	Percent. Answering Correctly		
		Cal.	Q'land	E. & W.
Section 1				
1	Add: 4 row t.u. with carrying figures	85%	91%	93%
3	Multiply: h.t.u. by 6	67%	80%	76%
4	Divide: h.t.u. by 7	59%	91%	78%
5	Add: 3 row ft. in.	7%	65%	61%	
10	Subtract: gal. pt.	3%	65%	49%
12	Divide: lb. oz. by 9	1%	71%	57%
30	Multiply: hr. min. by 6	0%	33%	24%
Section 2				
2	John has 7 sweets and I have 3 times as many. How many have I?	85%	91%	84%
5	Underline the smallest fraction: $\frac{1}{3}$, $\frac{1}{10}$, $\frac{1}{8}$, $\frac{1}{6}$, $\frac{1}{4}$	66%	37%	59%
12	Father was 32 years old in 1953. When was he born?	42%	61%	46%
13	What is half of $9\frac{1}{2}$?	15%	27%	61%
15	Write down the number with eighteen hundreds and five units	13%	68%	37%
25	How many pints of water can be drawn from a 22-gallon barrel?	7%	53%	45%

*The numbers in the items quoted here have been changed slightly

Tables 9 and 10 illustrate further the slightly superior performance in arithmetic of the sample of children from Queensland, Australia, over their counterparts in this country; they also illustrate the markedly inferior performance of the sample of children from California, U.S.A., and it is interesting to examine some of the factors that may have contributed to this latter result.

It is, of course, necessary to appreciate that there are a number of fundamental differences in the educational systems of the two countries any or all of which may have had some part to play in the differences observed in this study.

(i) Curriculum Differences

In this country the State plays little part in determining the curriculum to be followed by its schools, but California, like other American States, prints its own textbooks and largely dictates what is taught grade by grade. The age at which the arithmetic appropriate to each of the seventy items of the test was taught can therefore be determined. One item appeared in the Californian textbook for Grade III (i.e. 8-9 years); nine in the book for Grade IV (9-10 years); thirty in the book for Grade V (10-11 years); and thirty in the book for Grade VI (11-12 years). Since some of the children in the age-range considered were still in Grade V they had not yet studied some of the items, and hence could not be expected to get them right. It is not possible to determine in a similar way the exact age at which the items were introduced to British children, since most of the textbooks available in this country do not state the age at which arithmetical processes should be taught. However, an analysis of the Beacon Arithmetic Books, which do suggest appropriate ages, reveals that only one item appears at the same age in both countries, thirty-seven items appearing one year earlier in the English textbook and thirty-two items two years earlier.

One other important difference in the curricula of the two countries should be noted. Here the basic arithmetic syllabus is completed at the age of eleven and children move on to secondary schools. American children continue their elementary, (primary) education through grades 7 and 8 and do not begin secondary school until the age of 14. "But," according to Professor Buswell, "while American children are finishing arithmetic in grades seven and eight, English children of corresponding age who enter their grammar schools are covering two years of high school mathematics. The issue is whether England puts too much pressure on pupils or whether California delays its program too much."

(ii) Grade Placement v. Age Grouping and Streaming

Most of the children attending schools in England and Wales are streamed for ability within each year group, and the brighter children are encouraged to proceed at a faster pace than those in the "B" and "C" streams. In addition, as has already been noted, it is not an uncommon practice for the more efficient teachers to be assigned to "A" streams and the relatively inexperienced ones to the "C" streams. In the American schools all children within each grade move at very much the same pace although those considered not to have covered the ground of one grade may repeat it for a further year instead of being promoted. It might appear, therefore, that the British system would necessarily result in a better average performance, but it must be remembered that the system of grade placement is also employed in Australia.

(iii) Age of beginning school

English and Welsh primary schools admit pupils at age 5 whereas the customary age for the first school grade in America is 6. However, over 97 per cent. of the pupils in the Californian sample had attended a kindergarten admitting five-year-olds. The differences in performance cannot, therefore, be attributed to length of schooling, although the way in which the early school years are used may be a contributory factor. Little, if any, formal teaching is carried out in American kindergartens whereas in this country some formal instruction in reading and number is given to many five-year olds.

(iv) Teacher training

The Californian certificates of teaching require a four year college course of which approximately four-fifths is liberal education and one-fifth professional education. Thus, in terms of academic training the Californian teachers have had more years of preparation than most primary school teachers in this country. Elementary (primary) school teachers in the United States, however, suffer peculiar limitations in the case of arithmetic, since the subject is usually not taught in high school nor is it part of the mathematics syllabus in college. Hence, except for short review courses sometimes provided by teachers' colleges, elementary teachers have comparatively limited advanced training in teaching the subject. As a result, the specialised study of arithmetic for most American teachers does not go far beyond what they received in the last year of the elementary school—which is not very much more advanced than that taught to eleven-year-old primary school children in this country. It seems fairly clear, therefore, that teachers in this country probably have a considerably better background in arithmetic than do teachers in California.

(v) Different Social Conditions

While it may be clear that the factors considered so far contribute in varying degrees to the observed differences it is necessary in reviewing the evidence to consider two other facts which might have an adverse effect on arithmetic attainment in California. In the first place there is a fairly high degree of mobility of the population in California compared with England and Wales, due in part to migratory agricultural workers, in part to the prevailing tendency to move frequently from one location to another and in part to the continuous and large influx of population from other parts of the country. In two of the larger schools included in the sample the change in school enrolment during the year amounted to more than 100 per cent. A second possible adverse factor is that of race and nationality. In two of the larger schools included, more than 98 per cent. of the pupils were negroes. The mean scores on the seventy items for these schools were 6.2 and 6.6 compared to the mean of 12.1 for all schools in the Californian sample. Also, judging as nearly as possible from the names of the children, there were 211 Mexicans in the sample who made a mean score of 9.3.

VI. Summary and Conclusions

A national survey of basic attainments was carried out in England and Wales in May, 1955. One of the age groups concerned consisted of children aged between 10 years 9 months and 11 years 8 months at the time of testing. The tests used were specially prepared tests of arithmetic, reading comprehension and non-verbal ability.

These tests were applied in November, 1955, to a comparable sample of children in Queensland, Australia.

A modified form of one of the tests of arithmetic was also administered in January, 1957, to a comparable sample of children in California, U.S.A.

The results of the surveys have been compared and the following is a summary of the major findings.

1. The performance of the English and Welsh children was found to be superior in both reading and non-verbal ability to that of the Queensland children. (The latter are regarded as average in these respects in relation to other Australian children, but as somewhat superior in arithmetic.)

2. No significant difference was found between the mean scores of the two groups in the tests of arithmetic.

3. A closer inspection of these results, however, revealed a number of interesting differences.
   (*a*) Few of the Queensland children had been introduced to decimals or fractions and were, therefore, at a disadvantage in attempting some of the items in the tests.
   (*b*) The Queensland children's performance was also poorer than that of the children in this country in the more complex problems. This is probably related to the difference in levels of reading comprehension.
   (*c*) The Queensland children were also shown to be at a disadvantage in the test which involved strict time limits and required a fast rate of working. This difference may perhaps be ascribed to the fact that many children in this country, as part of their preparation for the eleven plus examination, are practised with speeded tests.
   (*d*) Taking these factors into account the analysis showed that the Queensland children were markedly superior in mechanical arithmetic involving the application of the four rules, conversion units, etc.

4. The standard deviation or "spread" of the scores obtained by children in this country was larger in all the tests. This may in part reflect the effects of "streaming." It is sometimes alleged that one of the results of streaming is that the brighter children make more rapid progress than they would otherwise achieve, and also that the duller children, when assigned to a "C" stream, become more backward, partly as a result of poor morale, and partly because there is a tendency in some schools for the more experienced teachers to be put in charge of the abler streams. It should be noted, however, that as far as reading is concerned, there were fewer "poor" readers found in the English and Welsh sample than in the Queensland group.

5. The comparisons with the children in California were based on one of the arithmetic tests only. (This was modified somewhat because some of the items—especially those involving money sums—were not applicable in American schools.) The children in Central California are regarded as about average in attainment in relation to children in the United States as a whole.

   The performance of the Californian children was found to be markedly inferior to that of children in England and Wales in both the mechanical and the problem arithmetic items.

Although children in California do not enter the elementary schools until 6+, the majority of the sample had in fact attended kindergarten schools. The main reasons for the different levels in performance are probably the fact that formal teaching tends to be introduced at an earlier age in this country, and the fact that there is a difference in the standards in the two systems. In the United States more limited objectives are formulated for children of primary school age, less emphasis being placed on rapid progress in mechanical arithmetic than is customary in this country.

6. It is interesting to note that Queensland and California both have the grade system of organisation as distinct from the form of organisation prevalent in this country whereby children are promoted by age and, in the majority of primary schools, streamed or grouped according to ability. Nevertheless, the results show that Queensland children tend to be superior, Californian children inferior to English and Welsh children in arithmetical attainments. This would suggest that the form of organisation is not the most important factor in determining levels of attainment. The most important factor would seem to be the definition of standards and objectives and the amount of time and effort that it is considered desirable to devote to any particular aspect of the curriculum. No doubt, there are also differences, as the comparison between California and this country suggested, between the kinds of training and experience that teachers enjoy and their consequent effectiveness, which may have a bearing on the standards achieved by their pupils.

Abstract

FACTORS RELATED TO READING DISABILITIES IN THE FIRST GRADE OF THE ELEMENTARY SCHOOL

Doctor Thesis, University of Stockholm, 1958. *Acta Universitatis Stockholmiensis, Stockholm Studies in Educational Psychology* 2. *Stockholm: Almquist and Wiksell.*

by EVE MALMQUIST

THIS study deals in a most thorough and comprehensive way with the early stages of learning to read and the causes of reading disability in children of 7-8 in Swedish schools, that is, at the end of their first school year. It grew out of the author's experiences as reading teacher and psychologist in the municipal primary schools of Malmo and from an analysis of the results obtained in over 350 (cited in Bibliography) Danish, English, German and Swedish studies.

Purposes

The main purposes of the investigation were limited to the following: (i) to examine critically the concepts of reading disability so far used, and to suggest an operational definition capable of exact description; (ii) to produce the necessary measuring instruments for selecting children in need of remedial treatment; (iii) to clarify the relationship of certain factors with reading disabilities at the end of the first grade; (iv) to analyse the types of reading and spelling errors made at this stage; and (v) to see how far such disabilities tend to persist throughout the primary school course.

Method

To answer these questions the author constructed and standardised a battery of five tests covering different aspects of oral and silent reading, two spelling tests and five tests of visual perception. These instruments were of satisfactory reliability and validity and compare favourably with similar tests in use elsewhere. The battery of visual perception tests which includes figures, letters, digits, geometric, nonsense and various other configurations is particularly thorough and ingenious. Malmquist also developed rating scales for those variables which are not readily susceptible to quantitative measurement, e.g. personality traits, and collected other relevant data from interviews and various official sources.

Definition of 'Good' and 'Poor' Readers

Examination of the many concepts of reading disability put forward by medical and educational writers revealed that for the most part these are based upon unverified causal assumptions or are too vague. The author defined his concepts in meaningful statistical terms by reference to his instruments of measurement. Briefly, 'poor' readers are those children whose reading index (combined scores on his five nationally normed reading tests) falls at or below one sigma from the mean. 'Good' readers correspondingly have

scores at or above one sigma from the mean, and ' Medium ' readers are those who lie within the limits of plus or minus one sigma from the mean. A sub-group of children with ' special ' reading disability is also distinguished, consisting of pupils whose reading index is one sigma away from their intelligence quotient.

Sampling

For the main investigation random samples (homogeneous in respect of age and residence) were taken from all first-grade pupils in two school districts: Kristianstad and Linkoping. These numbered 399 children aged seven to eight who were divided into groups of ' poor ', ' medium ' and ' good ' readers according to the criteria, and assessed with regard to the more significant physical, intellectual, social and developmental factors which have been adduced as associated with reading ability.

Each factor was studied separately in its relation to reading disabilities but, as a supplement to the more traditional methods of investigation, an attempt was made to study also the interactions of several variables at a time by using the method of analysis of variance of higher order (multifactorial design).

Results

(i) *Ability*

Statistically, significant differences existed between poor and good readers with regard to intelligence (as measured by the Terman-Merrill individual test scale).

(ii) *Personality traits*

All but one (the ' need for social contacts ') of the personality traits considered significantly differentiated good from poor readers. Thus, besides being generally less intelligent, poor readers tend to have less self-confidence, less persistence and concentration, less ability to make contacts, and to be less stable and more inclined to submit to the leadership of others than are children of good reading ability.

(iii) *Socio-economic variables*

Of the socio-economic and socio-cultural factors studied, parental social status, joint taxable income, and education above the elementary school level, as well as the number of rooms and books in the home were all strongly associated with good reading in children. ' Broken ' homes, family size, and a child's ordinal position in it, were not shown to be significantly related to reading ability.

(iv) *Pre-school history*

Consideration of the children's pre-school development revealed that premature birth and an exceptionally low birth weight were more common among poor readers, but that birth injuries, speech defects, severe diseases and non-attendance at a nursery school were not significantly associated with either group.

(v) *Physical and physiological variables*

Left-handedness and defects of hearing and vision were not found to be related to poor reading ability. However, mean differences between poor and good readers in the scores on the visual perception tests were significant in every case. It therefore seems likely that further work on this test battery would be useful diagnostically.

(vi) *Teacher variables*

A significant relationship between the teacher's age, number of years' service and the pupils' achievements in reading was revealed, but the teacher's knowledge and efficiency as indicated by his training college diploma and final teaching mark bore little relation to them.

(vii) *Absence*

The hypothesis that poor readers are absent during the first school year to a greater extent than medium and good readers was not verified in this investigation.

(viii) *Sex differences*

An analysis of sex differences revealed that the girls obtained somewhat better results than the boys in spite of their lower average ability as measured by the Terman-Merrill. No significant differences were obtained in most of the other variables except that among girls there were fewer premature births and teachers rated them higher in respect of the personality characteristics which were shown to differentiate strongly between poor and good readers, i.e. persistence, concentration and emotional stability.

(ix) *Special reading disability*

Children with 'special' reading disabilities in the sense that their reading ability was very low (below one sigma on the distribution curve for the reading tests employed) but whose intelligence was normal or supernormal (I.Q. 100 or above on the Terman-Merrill scale) occurred in 4.8 per cent. of the whole combined sample, and when the line of delimitation for normal intelligence was drawn at an I.Q. of 90 this percentage was increased to 8.5. Pupils whose reading ability deviated significantly (more than one sigma) in a negative direction from general intelligence occurred in 20.8 per cent. of the whole combined sample.

Using the case-analysis approach, a separate study of children with reading difficulties and an I.Q. of at least 90 (34 cases) was also made. This showed that reading disabilities at first-grade level are never isolated effects, for all such pupils deviated negatively, in a very marked manner, from the mean for the total population investigated with regard to the major factors studied.

(x) *Persistence of difficulties*

In order to get an idea of the degree of consistency of reading disabilities and also to ascertain the predictive value of the tests used, a follow-up study of all but one of the original samples of children was made when they were in grade four. Correlations

between the results of reading tests given in the first grade and marks obtained by the children in the fourth grade were all above .6. Furthermore, with only one exception, all pupils designated as poor readers at the beginning of the investigation remained at a level of achievement below the average of that of their school fellows of the same age in the populations investigated.

(xi) **" Word Blindness "**

One other finding is of importance since it appears to dispose of the misleading notion of ' word blindness '. Except for the omission of vowels in spelling, all the various alleged symptoms of this ' disease ' entity were found in readers at all levels although to a greater extent among subnormal readers. The author therefore concludes that differences between poor and good readers are of a quantitative rather than a qualitative nature.

Conclusion

This study combines scientific objectivity with such a wide scope that its contributions to knowledge are many and varied. Its wide representative sampling from a normal population is a departure from the usual method adopted by research workers of selecting children from reading clinics, remedial reading classes, and child guidance clinics. In this way a better opportunity is afforded for studying representative groups of poor readers and children with ' special ' reading disabilities. The application of operational methods for the differentiation of various groups of readers in conjunction with these sampling principles also makes it possible to obtain an unequivocal and clear definition of the incidence of reading disabilities in any population.

Besides its methodological value this study makes a direct contribution to an important area of applied educational psychology and child development. It strongly supports the hypothesis of multiple causation and the notion of reading disability as a phenomenon within the normal range of variation.

J.M.M.
W.D.W.

A SELECTED AND ANNOTATED BIBLIOGRAPHY

J. B. Biggs

Explanatory Notes

The inclusion of a particular book or article in the bibliography which follows indicates that, in the compiler's view, it is worthy of consideration by teachers. The fact that a given book does not appear does not necessarily imply that it has been rejected, but simply that it may have escaped the notice of the compiler.

The publications listed below fall into three categories. The first is mainly concerned with the teaching of arithmetic in primary schools, although secondary education is covered to a limited extent; the second deals with books which put forward somewhat less orthodox approaches to the teaching of number, and with general books from overseas, notably the United States; the third covers reports of research and books on the psychology of arithmetic.

A caution should be added with regard to those books in Section II which advocate new methods of teaching. While each method has been tried out in at least one school, there has not been, as yet, any controlled research definitely establishing its worth, or otherwise. This is not to say that new methods are of no immediate value, but simply that one must exercise due care in accepting their authors' claims. While awaiting objective validation, it is, nevertheless, refreshing—and important—to be at least aware of different approaches to the study and teaching of arithmetic.

The asterisks (*) beside each title give an indication of the level of difficulty in terms of the background knowledge of statistical or psychological theory, etc., presupposed by the author.

*—Easy. **—Moderate. ***—Difficult.

IMPROVEMENTS ON CURRENT METHODS

***Arithmetic in Primary Schools.** Association of Teachers in Colleges and Departments of Education. Longmans, Green & Co., London, 1954. Sixth impression, 76 pp.

This is a report on the teaching of mathematics in primary schools in England. The purpose of the book is to supplement American reports. The writers try to show how the content of the subject, as distinct from method, may be enriched; how the content and suggested standards may be evaluated; and how the curriculum may be evaluated in the light of the A.T.C.D.E. Committee's observations, and available research. One interesting view the committee puts forward is that ' spatial knowledge ', or informal geometry, should be taught from the earliest stages, replacing a proportion of the purely arithmetical curriculum.

***Romance in Arithmetic.** M. E. Bowman. U.L.P., 1950. 160 pp.

A book containing illustrations and amusing examples designed to relieve the drudgery involved in mastering the basic number facts and processes.

***Traditional Number Rhymes and Games.** F. Doreen Gullen. Scottish Council for Research in Education, No. XXXII. U.L.P., 1950. 143 pp.

This collection of rhymes and jingles was assembled at the request of the Scottish Council's panel on early Number teaching. The aim was two-fold: to provide infant teachers with informal number material, and to try to recover the dying heritage of Scottish number rhyme. While many of the rhymes are in a rich Scots dialect, there are others which should be very suitable for the infant teacher in English schools.

****The Principles of Arithmetic.** A. Keith and J. Robertson. Blackie, London, 1949. 278 pp.

This book originates from lectures given by the authors to Training College students, and is directed in the main to Junior and Secondary Modern teachers. It is the authors' belief that " the teacher should have a knowledge of arithmetic far in excess of the standard to which he will be required to teach the subject " and the tone of the book is in keeping with this end. The reference to past research and the close analyses of methods of teaching the four rules are extremely useful.

***The Teaching of Arithmetic in Primary Schools.** K. Lovell and C. H. J. Smith. U.L.P., 1953. 128 pp.

This is a practical book addressed largely to Training College students and newly qualified teachers. It covers most aspects of the Infant and Junior curriculum. As in the first book mentioned, these authors recommend that more attention should be paid to ' spatial knowledge'. Other recommendations include such practical advice as how to get the best out of chanting tables, the equal-additions method in subtraction, and, in general, as much pre-, extra and intra-school number experience as possible.

***The Teaching of Mathematics in Post-Primary Schools.** J. H. Murdoch. Studies in Education, No. 12, N.Z.C.E.R. O.U.P., 1950. 183 pp.

Mr. Murdoch writes primarily for New Zealand Secondary teachers but nevertheless, a great deal of what he has to say about the general problems of secondary mathematics—particularly as it concerns Secondary Modern teachers—is erudite, practically sound and highly readable. Newcomers to the teaching profession will find Mr. Murdoch's maxims most valuable.

***Fundamental Number Teaching.** R. K. and M. I. R. POLKINGHORNE. Harrap, 1955. 196 pp.

This book was written in response to the requests of 2,000 teachers for advice from *The Teachers' World*, and has thus a directly practical aim. The authors start from the very first number experiences and work right through the Primary School curriculum. With Schonell, the authors recommend that basic addition facts should be learned systematically, in the same way as, sooner or later, the child must master the multiplication tables. Chapters especially worthy of note are those on place value and zero, and on methods of subtraction.

***Diagnosis and Remedial Teaching in Arithmetic.** F. J. and F. E. SCHONELL. Oliver and Boyd, London, 1957. 195 pp.

This book supersedes the authors' previous book, *Diagnosis of Individual Differences in Arithmetic*, and covers more ground in the light of the most recent research. In the first part of the book the authors consider the factors making for good achievement in arithmetic, what research has to say on arithmetic teaching, and the various causes of backwardness. Then follows a discussion on diagnosis of difficulties in arithmetic, and the way in which the Schonell Diagnostic Arithmetic Tests can assist diagnosis, and finally, a programme of remedial teaching on the basis of the Schonell test results. Thus, the book is of great importance to teachers in two ways: it outlines the most common difficulties encountered by pupils in learning the fundamentals of arithmetic; and it suggests methods for diagnosing and remedying the difficulties of individual pupils.

***Number.** THYRA SMITH, BLACKWELL. Oxford, 1954. 188 pp.

***A Background of Primary School Mathematics.** L. D. ADAMS. O.U.P., 1953. 189 pp.

These two books are written by retired H.M.I.'s both of whom have had considerable experience as teachers and inspectors. On the basis of this experience, they offer eminently readable impressions and suggestions of interest to Primary teachers and parents. The sections devoted to suggestions for widening the scope of informal number work are of particular value.

***The Teaching of Arabic Numerals.** G. G. N. WRIGHT. S.C.R.E., No. XXXVI. U.L.P., 1954. 45 pp.

This is an interesting history and analysis of the construction and form of Arabic numerals. The author analyses ambiguities and mistakes commonly made by adults and infants in the formation of numbers, and suggests a simplified and less ambiguous technique for teaching the writing of number. It should be of interest especially to the Infant Teacher, and also to those who teach handwriting and art.

*****The Mathematical Association** from time to time issue reports (published by G. Bell & Sons) specifically addressed to teachers and almost entirely concerned with the teaching of mathematics. These reports are very detailed and thorough, and should afford considerable assistance to the teacher in the planning and presentation of his subject. The journal of the Association is *The Mathematical Gazette*, issued four times a year, and is available to non-members. The post-war publications are as follows:

INTERIM REPORT ON THE TEACHING OF MATHEMATICS IN SECONDARY MODERN SCHOOLS; No. 29, 1949. (Out of print.)

REPORT ON THE TEACHING OF MATHEMATICS IN SECONDARY TECHNICAL SCHOOLS; No. 30, 1949. (Out of print.)

Report on the Training of Teachers of Mathematics in Technical Schools and Colleges; No. 31, 1949.
Report on the Teaching of Trigonometry in Schools; No. 32, 1950.
Report on the Teaching of Calculus in Schools; No. 33, 1951.
Report on the Teaching of Higher Geometry in Schools; No. 34, 1953.
The Use of Visual Methods in Teaching Mathematics; No. 35, 1954.
A List of Books Suitable for School Libraries; No. 36, 1954.
Report on the Teaching of Mathematics in Technical Colleges; No. 37, 1955.
Report on the Teaching of Mathematics in Primary Schools; No. 38, 1955.

II. THE TEACHING OF ARITHMETIC—*Less Orthodox Approaches.*

***Counting and Arithmetic in the Infants School.** W. G. Bass and O. S. Dowty. Harrap, London, 1956.

This is an account of a method of teaching arithmetic, which has been tried out in an infants' school. The book deals chiefly with infant work, describing in detail the method and apparatus used. The method is based on the proposition that " Arithmetic becomes largely an exercise in translation—from formal activities and words into symbols and signs . . ." The means by which this ' translation ' is carried out is based entirely on counting. Numbers are shown to be counting symbols and the different ways of counting are related to the four rules—each process being a logical development of the preceding one. The novelty of the approach to the subject should be of great interest to all infant teachers.

***Arithmetic in Action.** E. Brideoake and I. D. Groves. U.L.P., sixth impression, 1954. 128 pp.

This is the report of an experiment using the ' realistic ' approach. Life situations are replicated on a smaller scale for use in the Infant Department, and constitute the entire formal number experience of the infants concerned. Full details of the method and the easily constructed apparatus are given. The main claim of the authors is that with this approach the children develop a far more favourable attitude to arithmetic and without any lowering of the academic standards eventually attained.

****How to make Arithmetic Meaningful.** L. J. Brueckner and F. E. Grossnickle. John C. Winston Co., Philadelphia, 1947. 513 pp.

" The major objectives of the modern arithmetic program are: (1) to develop in the learner the ability to perform the various number operations skilfully and with understanding; and (2) to provide a rich variety of experiences which will assure the ability of the pupil to apply quantitative procedures effectively in social situations in life outside the school " (p. 1). Pages 2-513 of this book then proceed to develop the second of these themes. A great deal of research is reported in the book, but it is almost entirely related to U.S. conditions and hence needs intelligent adaptation to this country. Chapters of interest are those on ' The Modern Arithmetic Program ' and ' The Changing Arithmetic Curriculum ', Chapters V to IX, covering the first steps in arithmetic to decimal fractions, might be immediately useful. The last four chapters cover such topics as testing, remedial work, problem solving and apparatus.

****The Elementary School Subjects** (Part III: Arithmetic). Luella Cole. Rinehart & Co., Inc., New York, 1946, pp. 321-448.

This publication, like that of Brueckner and Grossnickle above, demonstrates the difference between British and American attitudes, particularly to the content of the syllabus. There are topics treated here, however, which are not usually dealt with in British texts, and these are of definite relevance—for example, the chapters on the analysis of difficulties with vulgar fractions and the solution of verbal problems.

****Numbers in Colour.** G. CUISENAIRE and C. GATTEGNO. Heinemann, London, 1954. 39 pp.

This book explains the method (originating in Belgium and in growing use elsewhere of teaching arithmetic to primary school children by means of Cuisenaire's carefully constructed system of coloured rods. Taking 1 cm. as a convenient unit length of rod, the remaining numbers are treated in ' factor families ', each family (e.g. numbers having a factor of 3) having an allotted range of colour. The idea is that by this means, the child will associate an extrinsic property (colour) with an intrinsic and necessary property of number, and thus, by experiment and active manipulation, he will acquire a ready understanding of number concepts and the four rules of number. This apparatus has been compared with the American Stern material, the main differences being that the latter makes no use of colour relations, and is more expensive and complicated.

****Teaching the New Arithmetic.** G. M. WILSON, M. B. STONE and C. O. DALRYMPLE. McGraw Hill Book Co., New York, 1939.

This book reflects clearly the essentially pragmatic approach which characterises much of the education in the U.S. More specifically, these authors adopt the " more rational view of arithmetic as a functional service tool. . . . Common adult usage should be the limit of any arithmetic undertaken for drill mastery. Work beyond this . . . if recreational . . . can be left to choice . . . if vocational . . . can be left to learning on the job by the few who need it . . . or if encyclopaedic, it can be left in the encyclopaedia." The authors, on the basis of much sound research, outline those topics in mathematics which are likely to be of practical use in adult life; those which are not are to be rigorously excluded from the curriculum. While many in this country might not be prepared to go as far as these authors, particularly in regard to the implications for technical education, this is an interesting and refreshing book, the chapters on problem work especially so.

III. GENERAL RESEARCH AND THE PSYCHOLOGY OF ARITHMETIC.

***Problems in Mathematical Education.** H. DYER, R. KALIN and F. LORD. Educational Testing Service, Princeton, 1956. 50 pp.

Following expressions of public dissatisfaction with the teaching of mathematics, the Carnegie Corporation asked the Educational Testing Service to describe the kinds of research activities likely to lead to better courses and teaching in mathematics. The result is this booklet, summarising past work in a bibliography consisting of 250 references, with general remarks as to the present implications and the suggested trend of future research. There are ruthless criticisms of training requirements, current attitudes to the subject and ' teacher readiness ', some of which might well apply to conditions in this country also.

****The Teaching of Arithmetic.** NATIONAL SOCIETY FOR THE STUDY OF EDUCATION, 50th Year Book, Part II. University of Chicago Press, 1951. 308 pp.

This yearbook is a summary of the most recent and important research projects on the teaching of arithmetic. Naturally, the context is that of the American educational system, but much of what is said is relevant to English schools. The authors stress ' inductive ' and ' relational ' methods of teaching—an emphasis based more on theoretical grounds than on strictly experimental research. Chapters of more direct practical interest are those by Swenson Spitzer and Horn. There are other valuable chapters bearing on wider aspects of number teaching—for example, the mathematical background needed by teachers of arithmetic and suggestions for research.

***Studies in Arithmetic.** Volume I, S.C.R.E., No. XIII. U.L.P., 1939. 144 pp.

This is the first progress report of the Scottish Council Sub-Committee on arithmetic. While it is true that much of the research is restricted in applicability and partly superseded by later investigations, nevertheless there is a good deal of work that is of continuing value and importance, particularly to the primary school teacher. Among the topics covered are: the time allotted to arithmetic, the relative difficulty of basic number facts (see also Publication Nos. XXVIII and XXXVII) and adding upward versus adding downward.

***Studies in Arithmetic.** Volume II, S.C.R.E., No. XVII. U.L.P., 1941. 218 pp.

This is the second collection of number researches published by the S.C.R.E. The samples were restricted to Scots children, and the research was undertaken over 15 years ago, but the following chapters especially are recommended as being important for teachers: the relative merits of methods of teaching subtraction (all the evidence points to the superiority of the equal-additions method); zero and its attendant difficulties; analysis or errors in fractions; and finally, an intricate report on placement of topics in arithmetic, in which the findings of the Washburne Committee of Seven are criticised.

***The Child's Conception of Number.** JEAN PIAGET, ROUTLEDGE and KEGAN PAUL. London, 1952. 248 pp.

This book is based on a series of experiments or test situations planned to investigate the genesis of number concepts from early infancy to adolescence, when the child should be ' arithmetically mature '. Piaget's main hypothesis, which he holds to be proven by his evidence, is that the construction of number concepts is concomitant with the development of logical thought, and that therefore logical and arithmetic operations constitute a single, psychologically natural, system. He sees the notion of ' constancy ' or ' conservation ' as being the primary basis of the psychology of arithmetic, as it is in the case of the physical sciences.

In his investigation of cardinal and ordinal quantities, he discovers three stages in all the operations considered. The first stage is based solely on perceptual judgments of ' greater ', ' equal ', ' smaller '; so that the four-year-old child does not readily perceive that the volume of water, for instance, remains constant when poured from a smaller to a larger container. The child then passes through an intermediate or transitional stage in which the visual percept may or may not be more powerful than the principle of constancy. Finally, around the age of seven years, the average child suddenly grasps the ' primacy of conservation '. The implications of this for the primary school teacher are evident. It is, for example, clear that until the child has reached this third stage he has little hope of comprehending the essential meaning of ordinary addition and subtraction, or even of measurement itself.

Criticism has been levelled at the possible lack of generality of Piaget's findings, at the ' loading ' of the outcome of his tests by the form of his questions. No-one, however, has seriously questioned his notion of sequence in development. The book represents an important contribution to psychological theory and may serve as a guide for the mathematics teacher who is prepared to translate its findings into practical terms.

For a useful summary of later developments in Piaget's work, in which number thinking is dealt with alongside other cognitive and affective aspects, the following is recommended:

***Berlyne, D. E.,** " Recent development in Piaget's work." *British Journal of Educational Psychology*, Volume XXVII, February, 1957, pp. 1-12.

****The Early Development of Number Concepts.** S.C.R.E., No. XX. U.L.P., 1942. 63 pp.

This is the report of a survey conducted by the Scottish Council amongst Scottish Infant schools on the extent to which children have acquired number knowledge—ability to count, do simple sums, etc.—before they enter Infant school. There are also smaller sections on frequency of use, in Infant schools, of number games, pictures and finger counting.

*****Addition and Subtraction Facts and Processes.** S.C.R.E., No. XXVIII. U.L.P. 66 pp.

This report is a sequel to the preceding one, and concerns the child's knowledge of number on his transfer to the Junior school. 1,315 children, from 47 Infant schools, were tested. Although specifically related to the Scottish system, the results are certainly suggestive and important for English teachers. It is seen from the tables published that the relative difficulty of items tend to remain constant throughout the school year; and that, generally, as the size of digits increases from 0 to 9, so does the difficulty of the item. Numbers containing zero, however, are difficult at any level and tend to show less consistent gradation. By obtaining the same sorts of data about processes, it becomes possible to construct tests of pre-arranged difficulty and with a predictable distribution of marks.

*****Patterns of Error in the Addition Number Facts.** James M. Thyne. S.C.R.E., No. XXXVII. U.L.P., 1954. 275 pp.

Using the same data as the report on *Addition and Subtraction Facts and Processes* (S.C.R.E., No. XXVIII), Thyne takes the research an important step further by enquiring into the *nature* of error—the psychological processes involved—as distinct from mere frequency of kinds of error. The author makes the point that when the teacher understands " answer-producing " mental activities, he has much greater hope of remedying mental processes productive of error. The author proposes a tentative psychological theory of meaning, closely analyses the errors obtained in the light of this, and then suggests practical implications which follow for the teacher.

TEST ANNOUNCEMENTS

Two distinct series of tests are constructed at the Foundation. They are:

" Closed " Series

Standardised tests for allocation at 11+ and for transfer at 13+ are produced each year. A catalogue giving information about these tests will be sent to local education authorities on request.

" Open " Series

Tests in this series have been prepared specially for the use of the teacher in the classroom. Details and prices are given in the catalogue, " Educational Guidance in Schools: Standardised Tests for the Use of Teachers," obtainable from the Foundation.

In addition to the above tests constructed by the Foundation's staff, a large number of tests and test materials are distributed through the Foundation's *Test Agency* on behalf of the Psychological Corporation of America, the Australian Council for Educational Research and various authors in this country. A catalogue giving brief descriptions of the tests, their availability and prices, can be obtained from the Foundation.

The following tests have recently been made available through the Foundation's Test Agency:

Wing Standardised Tests of Musical Intelligence

These tests are designed to measure the children's mental capacity to cope with music. The results can be expressed in terms of gradings according to age. These are of considerable help in giving some guidance in the following directions:

(1) Picking out musically bright children who can profitably study a musical instrument, e.g. for the school orchestra.
(2) Selecting children who are capable of studying music up to G.C.E. levels, or students who might profitably follow an Advanced, Ordinary or basic course at a Training College.
(3) Dividing a group of school children, or of students, into ' sets ' for the study of music.

The test consists of a number of standardised ear tests which are played on a piano and then recorded on tape. The tests include: Analysis of chords; Pitch discrimination; Memory for pitch; Harmony appreciation; Intensity appreciation; Rhythm appreciation; Phrasing appreciation. They are very easy to administer and are suited to a wide range of candidates of from 8 years to good adult status. Norms are based on many thousands of cases. The playing time of the test is about 45 to 50 minutes—say about an hour for testing students or two periods of about 40 minutes for testing children. Further details and prices are available from the Foundation.

A Battery of Diagnostic Performance Tests

With the publication of the book " Diagnostic Performance Tests " by Boris Semeonoff and E. L. Trist (Tavistock Publications, Limited, 1958) the following tests have now been made available through the Foundation's Test Agency:

1. The Semeonoff-Vigotsky Test.
2. The Trist-Hargreaves Test.
3. The Trist-Misselbrook Test.
4. The Carl Hollow Square.
5. The Passalong Test.

Further details and prices of these tests, and copies of the book, can be obtained from the Test Agency, 79, Wimpole Street, W.1.

©

National Foundation for Educational Research in England and Wales.
Printed by King, Thorne & Stace Limited, 1-13 Jubilee Street, Brighton, 1,
and published by Newnes Educational Publishing Co., Ltd., Tower House, Southampton Street, Strand, London, W.C.2.,
for The National Foundation for Educational Research in England and Wales, 79 Wimpole Street, London, W.1.

Publications of the

NATIONAL FOUNDATION FOR EDUCATIONAL RESEARCH

BOOKS

1. A List of Researches in Education and Educational Psychology. 1918-48. By A. M. Blackwell, M.A., B.Sc., £1 1s. 0d.

2. The Allocation of Primary School Leavers to Courses of Secondary Education. First Interim Report. By A. F. Watts, M.A., D.Lit., and Patrick Slater, M.A., 7s. 6d.

3. A Survey of Rewards and Punishments in Schools. Based on Researches carried out by M. E. Highfield, M.A., Ph.D., and A. Pinsent, M.B.E., M.A., B.Sc., £2 2s. 0d.

4. The Educability of Cerebral Palsied Children. By M. I. Dunsdon, M.A., F.B.Ps.S., £1 1s. 0d.

5. A Second List of Researches in Education and Educational Psychology, 1949-51. By A. M. Blackwell, M.A., B.Sc., £1 1s. 0d.

6. Secondary School Entrance Examinations. By A. F. Watts and others, 6s. 6d. net. 3s. 6d. (limp covers).

7. Lists of Researches in Education and Educational Psychology. Supplement I. 1952-53. By A. M. Blackwell, M.A., B.Sc., 10s. 6d.

8. Pupils' School Records. By Alice S. Walker, B.Sc., £1 1s. 0d.

9. Lists of Researches in Education and Educational Psychology. Supplement II. 1954-55. By A. M. Blackwell, M.A., B.Sc., 15s.

10. Admission to Grammar Schools. Third Interim Report on the Allocation of Primary School Leavers to Courses of Secondary Education. By Alfred Yates, B.A., B.Ed., and D. A. Pidgeon, B.Sc., £1 5s. 0d.

11. Lists of Researches in Education and Educational Psychology. Supplement III. 1956-57. By A. M. Blackwell, M.A., B.Sc., 15s.

12. Reading in the Primary School. By Joyce M. Morris, B.A., Ph.D., with a Statistical Appendix by P. M. Grundy, B.A., Ph.D., 15s. net.

11-9-6
1-7-6
10-12-6

CONTENTS

PAGE

EDITORIAL 2

A NEW LOOK AT INTELLIGENCE TESTING. *P. E. Vernon* 3

TRANSFER AT 11 PLUS. *Alfred Yates and D. A. Pidgeon* 13

THE WISH TO LEARN—RESEARCH INTO MOTIVATION. *W. D. Wall* 23

TEACHING CHILDREN TO READ—I: THE RELATIVE EFFECTIVENESS OF DIFFERENT METHODS OF TEACHING READING. *J. M. Morris* 38

A COMPARATIVE STUDY OF BASIC ATTAINMENTS. *D. A. Pidgeon* 50

ABSTRACT—FACTORS RELATED TO READING DISABILITIES IN THE FIRST GRADE OF THE ELEMENTARY SCHOOL. *Eve Malmquist* 69

THE TEACHING OF ARITHMETIC—A SELECTED AND ANNOTATED BIBLIOGRAPHY. *J. B. Biggs* 73